2025, Volume 48, Issue 96, 26-61/ ISSN 2304-4306

E C O N O M Í A

revistas.pucp.edu.pe/economia

FONDO EDITORIAL

www.fondoeditorial.pucp.edu.pe

Determinants of the industrial location in Peru, 1963-1974

Gonzalo Rivera Mariscal [a]

[a]Ministry of Economy and Finance of Peru
✉ g.rivera@pucp.edu.pe

Abstract
The present study responds to the following question: What factors determined the location of manufacturing industries in Peruvian regions for the years 1963 and 1974? Using the data of the Economic Censuses conducted in 1963 and 1974 by the Peruvian government, we evaluate the factors that influenced the location decisions. For this aim, we apply the methodology proposed by Midelfart-Knarvik et al. (2000, 2001), which integrates in a model the factors that the Heckscher-Ohlin (H-O) and the New Economic Geography (NEG) theories consider important to explain industrial location decisions. Among them, they consider the influence of the regional endowment of resources and the intensity of their use in industries (H-O theory), as well as the influence of the market potential and the backward or forward linkages between industries or the economies of scale in industries (NEG theory). Our findings indicate that for this period of analysis in Peru, the factors related to agricultural endowment, electrical energy, financial capital (components related to the H-O theory), backward linkages and the economies of scale (components related to the NEG theory) were influential in determining the industrial location decisions of the manufacturing sector across regions. The results also indicate that the two components associated with the NEG theory have the highest weighted impact on manufacturing location decisions. Another relevant aspect is that our findings allow us to partially understand the agglomeration of industries in some regions, particularly in the capital of the country, Lima.

Article History: Received: 2 February 2025 / Revised: 16 April 2025 / Accepted: 11 August 2025
Keywords: Industrial location, regional economics, factor endowments, new economic geography, economic history of Peru
JEL Classification: N66, N96, O14, O18, R12

Acknowledgements

The opinions expressed in this paper are those of the author and do not represent or compromise the position of the institution with which the author is affiliated.

https://doi.org/10.18800/economia.202502.002

1. Introduction

For decades, inequality has been a topic of interest in several aspects of social science. For example, the inequality of income has been explored in a very influential book by Piketty (2014), who describes how income inequality has evolved in different regions of the world. Some other important studies related to this topic are, for example, the ones developed by Milanovic (2016), Scheidel (2017), and Zucman (2019). Also, another type of inequality explored has been the unequal development among countries. For example, Acemoglu and Robinson (2012) elaborate an influential work that provides explanations about the factors that promote the development of nations. (Other relevant works include those by Acemoglu et al. (2002) and Spolaore and Wacziarg (2013)).

Nevertheless, another interesting topic arose in trying to understand how uneven geographical development impacts different ways of production in different localities. This particular topic takes impulse with the studies developed by Paul Krugman during the 1990s under the approach named *New Economic Geography* and has grown rapidly, perhaps for its analytical clarity (MacKinnon and Cumbers, 2019).

Considering that, the present research intends to provide more insights into the literature about the importance of geography in economics, by focusing on the particular case of Peru (an independent country from the Spanish Colony since 1821) and the determinants of the location of industries. Even though there have been different studies theorizing the factors that explain the industrial location (for example, Weber (1929), Lösch (1954) and Krugman (1991)) as well as some other empirical ones (for example, Kim (1995) and Rosés (2003)), there is still a relatively small number of studies that focus their attention in developing countries in Latin America.

One motivation to explore this topic using Peruvian regions as units of analysis is due to the results obtained by Seminario et al. (2019) and Seminario et al. (2020), in which the authors estimate the Gross Domestic Product (GDP) of the regions from the period 1795 until 2017. As we can observe in Figure 1, the regional GDP in Lima, the capital of the country, shows considerable development in comparison to other regions, with a remarkably different trend since 1930.

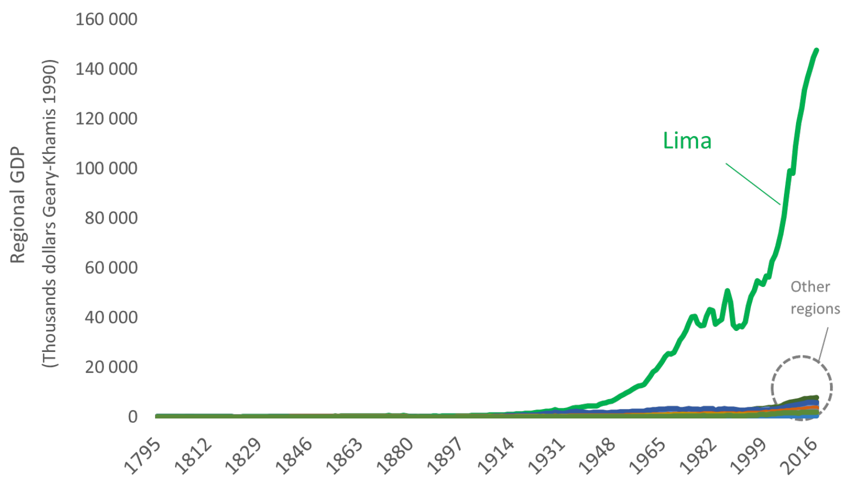
Studies describing the unequal development among Peruvian regions during different periods in the mid-twentieth century reassure us of the existence of this particular phenomenon. For example, Slater (1975) evaluates the unequal development in Peruvian regions between the years 1946 and 1955 by building an index that includes indicators of economic activity, educational participation, and social services, finding that Lima was at the top of the ranking with the highest score in all three indicators.

Additionally, Smith (1968), for the period 1961 and 1963, presents an evaluation of the regional income and living standards for all the Peruvian regions. By considering thirteen variables of regional wealth and development, he also finds that the region at the top of the ranking was Lima, with an urban population rate of around 90 % of the population in the region, in contrast to other regions whose rates vary between the range of 2 % and 52 %.

One influential factor could be associated with the development of industries. As Krugman and Venables (1995) have described theoretically, the process of concentration of industries in a region can create an unequal development among regions since the economy organizes spontaneously in an industrialized core and a deindustrialized periphery.

In this way, one initial point to understand this process is to realize how some regions were chosen for manufacturing industries as points of location. According to Vega-Centeno (1988), industrialization in Peru took a significant volume and dynamic since the 1950s, and therefore it makes it a period of interest. However, data available near this period, integrating all the industries in the country, is only available in the two Economic Censuses conducted by the Peruvian government in 1963 and 1974, which in turn define the periodization of the present study.

Figure 1. Historical GDP of Peruvian regions (1795-2017)



Elaborated with the data from Seminario et al. (2020)

Industrialization is widely considered a key driver of development; therefore, looking into its spatial determinants is relevant for understanding the causes of uneven development. For this reason, the present research tries to answer: What are the determinants that influenced industrial location decisions in Peru for the years 1963 and 1974?

It is expected that the results favor the New Economic Geography theory, which emphasizes that regional market conditions explain the dynamics of industrial location. This expectation contrasts with the Heckscher–Ohlin theory, which indicates that the resource endowment of a region can attract industries. The reason to consider that is because regions with valuable inputs for industries, such as abundant oil in Piura, or the rich wood inputs in Loreto, seem not to be key components to attract industries to their locations. For example, a comparison of the average of industrial establishments between 1963 and 1974 reveals that: While in the capital, Lima, 35 % of the total industrial establishments were placed there, in Piura and Loreto the share was 8 %, and 2 %, respectively (See Table 2).

The findings indicate that during the years 1963 and 1974, the factors with more weight influencing the industrial location were related to the New Economic Geography theory, in which the presence of economies of scale in the manufacturing industries was relevant. However, factors related to the Heckscher–Ohlin theory are also significant. The results of the present paper contribute to understanding the determinants of industrial location in a developing country in the region of Latin America, widening the current knowledge about the topic.

The present study is structured as follows: In section 2, we present the theoretical framework for the study, and in section 3, we review empirical studies that explain the location of industries in different countries, complemented also by some related studies about the Peruvian economy. In section 4, we describe the methodology applied for this research, and in section 5, we describe the data and the sources. We present the estimation and results in section 6. In section 7, we explain the results, considering also the Peruvian context during the period of the analysis. Finally, in section 8 we present the conclusion and policy implications of the study.

To illustrate how the regions are distributed across the country, we present a map with the current political regions of Peru (Figure 2), which differs slightly from the regions of our study since the two differences are the Ucayali region that was founded in 1980 (between 1963 and 1974 this area was part of the Loreto region), and the Callao region, which was created in 2003 (between 1963 and 1974 this area was part of the Lima region).

Figure 2. Regions of Peru

Source: Miranda et al. (2014)

2. Theoretical framework

To explain the factors that influence industrial location decisions, we find in the literature two main theories: The Heckscher–Ohlin (H–O) theory and the New Economic Geography (NEG) theory.

According to the Heckscher–Ohlin theory, in a scenario where there is perfect competition, homogeneous products, and non-increasing returns to scale, the industrial geographical distribution is determined by variations in natural endowments, technologies and factors that provide comparative advantages to the industries (Brülhart, 1998). Specifically, the theory proposes that in a scenario of trade among economies, the economy with a relatively abundant factor of production, in comparison with other economies, will tend to produce the goods that are intensive in the use of that factor (Krugman et al., 2018).

The New Economic Geography theory, on the other hand, focuses attention on the relevance of economies of scale, the input-output linkages between industries and transportation costs to explain the manufacturing concentration in a region or a country. According to Krugman (1991), in industries where economies of scale are present, firms will prefer locations with a considerable nearby demand since the transportation of the products through the markets is minimized.

Another factor that influences location decisions, according to the New Economic Geography theory, is the input–output linkages between industries. The backward linkage effects occur when in a region there is a high demand for intermediate goods, generating that producers of these goods will see the region as a more attractive place to locate their production (Krugman and Venables, 1995). Producers of these intermediate goods will prefer to be located near the firms that buy their products to minimize the transportation costs involved (Martínez-Galarraga, 2012). It is worth noting that some of the demand for manufactured goods will also come from the manufacturing sector itself (Krugman, 1991).

This creates a scenario where it occurs what Arthur (1990) and Pierson (2011) call a "positive feedback process", which means that the effects of an event begin to accumulate and provoke a self-reinforcing cycle. The positive feedback process appears because manufacturing industries will locate in areas with a large market, and as a result, the market will increase, attracting more industries and subsequently increasing the size of the market again, and the process goes on.

The forward linkage effects occur when a region with a great variety of intermediate goods attracts firms because there they can obtain inputs at a lower price and therefore reduce the cost of production of their final goods (Krugman, 1991; Krugman and Venables, 1995). Downstream industries will face lower prices if they locate in places with many upstream industries, as they save on transportation costs on their intermediate inputs (Venables, 1996). Ceteris paribus, it will also generate a positive feedback process as the market increases with the new industrial establishments because living and producing near a concentration of manufacturing is more desirable for industries, as it reduces the cost of purchasing goods provided by that central location (Krugman, 1991).

Additionally, this process will increase the demand for labor in the industrializing region with a subsequent increase in the real wages attracting more population at the expense of other regions, therefore generating an uneven development by increasing the concentration of the manufacturing activity in a few regions (Krugman, 1991; Krugman and Venables, 1995; Venables, 1996).

However, if the transportation costs fall low enough to reduce the importance of being placed in a location close to markets and suppliers, then the areas lagged in industrial production will be a focus of attention for businesses attracted by lower wage rates; additionally, it can produce mobilization of industries to move out from the core to the periphery, generating a convergence of wage rates (Krugman and Venables, 1995; Venables, 1996).

On the other hand, in a scenario with limited benefits from the economies of scale and significant expenses associated with transportation costs, the manufacturing sector can be obligated to be placed near the agricultural sector since they are the market that buys their products, generating, in that case, a dispersion of the manufacturing activity in different regions (Krugman, 1991). A scenario like this is usually presented in a pre-railroad and preindustrial society, but when the economies of scale increase due to mass production and the transportation costs are reduced, the tie to the land is no longer a feature, since establishing in places where there is a big market potential results more beneficial than near to the agricultural sector (Krugman, 1991).

In that way, the New Economic Geography identifies three different scenarios regarding transportation costs. A first scenario in which the transport costs are high, making the industries tied to land and generating a dispersion of manufacturing activity. A second scenario, in which the transport costs are lower, allows industries to locate or agglomerate near areas with high market potential and are no longer tied to land, since the cost of mobilizing the inputs necessary for their production is no longer a burden. And a third scenario, in which the transport costs are much lower than in the second scenario, in which the mobilization of the final goods to the markets or the intermediate goods to the production sites is no longer a burden, and the location can be subject to the cost of the labor force.

It is worth noting that both theories, the Heckscher-Ohlin and the New Economic Geography theories, do not exclude each other, in the sense that the location decisions of industries can be influenced by the factors both theories predict at the same time. When trade is perfectly free, the comparative advantage of areas (such as technology or endowment differences) determines the structure of production in each area; however, when transport costs or trade frictions are present, aspects like the market potential of the region, combined with comparative advantage, matter (Midelfart-Knarvik et al., 2001). Even, it is also possible that the original advantage of a place loses its relevance due to changes in technology (for example a change from a coal-based technology to an electric-based technology), but thanks to the expertise base and network of specialized inputs established there (New Economic Geography forces), the importance of the location persists (Crafts and Mulatu, 2005; Klein and Crafts, 2012), .

3. Literature review

Empirical studies

After the seminal studies of Krugman (1991) and Krugman and Venables (1995), there have been different authors trying to test the H-O and NEG theories since it is considered that the components related to these theories are capable of exerting influence at the same time in the location decision of industries.

For example, one first influential study was developed by Kim (1995) who analyzed if external economies, such as informational spillovers and availability of specialized factors ("Marshallian externalities"), the economies of scale (NEG components), and factor endowments of a region (H-O component) influenced the location trends of the United States manufacturing sector location between 1860 to 1997. While he finds no evidence that the externalities have an impact on the concentration of industries, he concludes that scale economies and resources are relevant, ;however, without providing the weight that each factor has in explaining this phenomenon.

Ellison and Glaeser (1999) also estimate the effects of the advantages provided by resource endowments on the location of manufacturing industries in the United States, finding that 20 % of the concentration of industries is due to firms choosing areas that provide them with cost advantages.

In another study, Kim (1999) explores the determinants of the geographic distribution of industries in the United States between 1880 and 1987, by considering the factor endowments and "Marshallian externalities". Even though he provides an argument to support the influence of natural advantages due to factor endowments, the findings are inconclusive concerning "Marshallian externalities".

Additionally, Rosés (2003) investigates the relevance of the Heckscher-Ohlin and New Economic Geography frameworks to explain the differences in industrialized regions in Spain from 1797 to 1910. The author concludes that the endowments from regions, as well as the home market effects (as a potential result of the market-size scale of economies and backward and forward linkages), explain the concentration of the manufacturing activity.

For the case of Latin America, there have also been some efforts to understand the determinants of industrial location, but the majority focus their attention around the beginning of the twenty-first century. For example, Ferreira et al. (2003) analyze the location of 12 manufacturing industries across Brazilian municipalities in 2001, assuming firms choose sites where expected benefits exceed a threshold. Using a spatial cost function at the firm-level, the study finds that infrastructure development that reduces transport costs attracts firms, while fiscal incentives have modest effects, agglomeration has no direct impact, and regional diversity slightly raises costs.

For the case of Colombia, Antošová and Arias Gómez (2024) examine manufacturing location in Colombia using cross-sectional data for 1,118 municipalities in 2005, and a panel data set with Colombian regions between 2011 and 2015. They find that manufacturing industries concentrate in densely populated areas with attractive markets, and that the unobserved effects in the panel estimation support the influence of factors not controlled in the regressions.

Toro González (2004) also explores the factors that explain the localization of manufacturing industries in eight main metropolitan areas of Colombia from 1990 to 1999, using a panel data estimation. The results show that labor costs negatively affect location decisions, while regional market size and especially sectoral agglomeration have positive effects.

In some studies, due to the methodology applied, it is not possible to test the influence of the

New Economic Geography forces, while in others, such as the studies for Latin America, even though they consider this theory, they omit factors related to the Heckscher–Ohlin framework. As a result, neither of them provides information about the relative weight of influencethat each component of the two theories simultaneously exerts on industrial location decisions.

One model that fills those gaps appears in Midelfart-Knarvik et al. (2000, 2001), who propose a methodology that allows for the simultaneous evaluation of the impact of the factors related to the H–O and NEG theories, as well as the relative influence of each factor in industrial location decisions. This model has been the base for different historical economic studies, which, in general, focus their attention on European countries or the United States during various periods between the mid-nineteenth and mid-twentieth centuries. These studies show that components of both the Heckscher–Ohlin and New Economic Geography theories can be simultaneously influential and significant in industrial location decisions.

One of the first studies to apply this methodology was conducted by Crafts and Mulatu (2005). The authors explore what explains the location of industries in Britain from 1871 to 1931, finding that factor endowments (H–O component), such as human capital, and the presence of market potential and scale economies (NEG components) in regions influence the location decisions of industries.

Additionally, Wolf (2007) analyses the forces that interacted during the period posterior to the Polish reunification (1925–1937), and concludes that not only the presence of forward linkages (NEG component) influenced industrial location decisions in regions, but also the availability of skilled labor force for skill-intense industries and the presence of patents for firms with high innovative activity (H–O components).

Klein and Crafts (2012) also investigate the factors that explain the formation of the "manufacturing belt" in the United States, which is an area of remarkably concentrated industries that prevailed during part of the nineteenth and twentieth centuries. The study focuses its attention on the years 1880–1920 and finds that the share of farmland in a state (H–O component) and the market potential, in conjunction with forward linkages and scale effects (NEG components), impact the location decisions of manufacturing in the United States.

Additionally, Martínez-Galarraga (2012) presents a study of Spain considering the period between 1856 and 1929. His results indicate that the market potential, working together with the scale effects (NEG component) and the agricultural abundance in a region (H–O component), had an influence on the location of industries..

Moreover, Ronsse and Rayp (2015) describe the factors that determine the location of industries in Belgium between the years 1896 and 1961. The authors find that the factor endowments of a region play a key role in the location decisions of industries, but the NEG forces, especially the interaction between market potential and economies of scale, have more weight, particularly from the twentieth century onward.

Nikolic (2018) also provides an analysis of the factors that impact the location of industries in the no longer existing Yugoslavia for the period between 1932 and 1939. His results indicate that the NEG theory is relevant to explain industrial location due to the influence of market potential together with input linkages, as well as the H–O theory, due to the effect of the availability of human capital, which becomes relevant for skilled labor intensity in firms.

Additionally, Missiaia (2019) investigates the factors that influenced the decisions of the location of the manufacturing industries in Italy for the period between 1871 and 1911, a period of the first industrialization of the country. The results of the study indicate that energy endowment and human capital (H–O forces), as well as aspects related to domestic market access (NEG forces), are relevant for determining industrial location.

This methodology has also been applied in some Latin American country cases. For Argentina, Sanguinetti and Volpe Martincus (2009) examine the determinants of manufacturing location between 1985 and 1994. They find that lower tariffs lead to industrial de-concentration, that input-intensive sectors cluster near established industrial bases, and that industries reliant on oil or minerals locate near resource-rich regions.

Badia-Miró and Yáñez (2015) also analyze the determinants of industrial location in Chile for the years 1911 to 1967. They find that before the Great Depression, endowment factors were relevant for industrial location decisions and the industries were dispersed; while after 1930, industries that exploit economies of scale concentrated in provinces with large market potential.

Regional economic and political economy studies in Peru

In the case of Peru, even though research that explores the determinants of the location of industries is limited, there are some papers that have contributed to the spatial economic studies in the country.

For instance, Herrera Catalán (2014) explores how agricultural transportation costs influence the industrial location of agricultural-related manufacturing industries in Peru by using data from 2013. Using a distance-based test, the author finds that these industries do not show strong localization patterns, and by using a conditional logit model, he also finds that firms localized at short (long) distances are explained by low (high) agricultural transport costs.

Additionally, one of the recurrent topics explored has been the idea of regional convergence in terms of GDP per capita among Peruvian regions. Examples of these studies are the ones developed by Gonzales de Olarte and Trelles (2004), Sutton et al. (2006), Delgado and Del Pozo (2011), Delgado and Rodríguez (2014), Delgado and Rodríguez (2015), and Palomino and Rodríguez (2019), which focus their analysis on the period between 1950 and 2010. Probably one of the most ambitious studies on this topic is the one developed by Seminario et al. (2019), who evaluate the convergence in Peruvian regions for the period between 1795 and 2017 and also provide estimations of the GDP for the different regions for the same period, by applying a variation of the method proposed by Geary and Stark (2002).

Additionally, Gonzales De Olarte et al. (2012) explore the formation of various important urban centers in the capital of Peru, Lima, during the first decade of the twenty-first century, by considering the employment density.

Regarding the economic context for the period of this study, it is worth mentioning that different studies have pointed their attention to the performance of the economy during the second half of the twentieth century when the Peruvian government followed their neighborhoods in the region in the process known as Import Substitution Industrialization (ISI), intending to promote more development of local industries and a less dependency on the importation of manufacturing products. For example, Abusada (1977) and Chumacero (2012) describe the policies implemented by the Peruvian government during the period 1930–1976 to promote industrialization in the country. On the other hand, Jiménez et al. (1999) and Thorp and Bertram (2020) analyze the performance of the industrial sectors in the economy. Additionally, Gonzales de Olarte (2016) analyzes how the Peruvian national market among sectors has changed in their level of integration since 1950.

Given that, the present research aims to complement the regional economic studies in Peru by adding an explanation of the factors that influenced industrial location decisions of the manufacturing sector in the country for the years 1963 and 1974. Moreover, it is important to mention that the present study also contributes to the literature by exploring the determinants of industrial location for a developing country in the region of Latin America.

4. Methodology

The methodology of the present study is based on the econometric specification proposed by Midelfart-Knarvik et al. (2000, 2001), but with the variation suggested by Klein and Crafts (2012). The advantage of the methodology is that it allows us to evaluate simultaneously the relative effect of the components identified by the Heckscher–Ohlin and the New Economic Geography theories as determinants of industrial location.

Consider first the following model proposed by Midelfart-Knarvik et al. (2000):

$$\ln(s_i^k) = \alpha \ln(pop_i) + \varphi \ln(man_i) + \sum_j \beta[j]\left(y[j]_i - \psi[j]\right)\left(z[j]^k - \phi[j]\right) + \xi_i^k$$

Where:

- s_i^k is the share that represents the output of industry k in region i in the total output of the industry k in Peru.
- pop_i is the share of Peru's population living in region i.
- man_i represents the share of the manufacturing sector output of region i in the total output of the Peruvian manufacturing sector.
- $y[j]_i$ is the level of the jth regional characteristic in region i.
- $z[j]^k$ is the industry k value of the industry characteristic associated with regional characteristic j.
- ξ_i^k is the error term.
- α, φ, $\beta[j]$, $\psi[j]$, and $\phi[j]$ are the coefficients.

The variables pop_i and man_i capture regional size effects; in this way, all else equal, it is expected that larger regions have a larger industrial share in any given industry. The remaining terms capture the interaction between the regional and industry characteristics. To provide a further specification of how to interpret the model, we replicate here the explanation presented by Midelfart-Knarvik et al. (2000):

Let us think about a particular characteristic, say $j = capital$, so $z[capital]^k$ is then the capital intensity of industry k and $y[capital]_i$ is the capital abundance of region i. The model can be interpreted as:

1. There exists an industry k with capital intensity $\phi[capital]$ whose decision of location is not affected by the capital abundance of the region.
2. There exists a level of capital abundance, $\psi[capital]$, in which the region i's share of each industry is independent of the capital intensity of the industry.
3. If $\beta[capital] > 0$, then industries with capital intensity greater than $\phi[capital]$ (therefore, $z[capital]^k > \phi[capital]$, or $z[capital]^k - \phi[capital] > 0$) will be drawn into regions with capital abundance greater than $\psi[capital]$ (since also $y[capital]_i - \psi[capital] > 0$), and out of regions with capital abundance less than $\psi[capital]$.

In this way, a key factor of the model is $\beta[j]$, which gives us information about the relevance of different regional characteristics in location decisions by taking into account the relative sensitivity of all industries to those characteristics.

To estimate the coefficients, the specification is expanded to provide the following estimating equation:

$$\ln(s_i^k) = c + \alpha \ln(pop_i) + \varphi \ln(man_i) + \sum_j \Big(\beta[j]\, y[j]_i\, z[j]^k - \beta[j]\psi[j]\, z[j]^k - \beta[j]\phi[j]\, y[j]_i\Big) + \xi_i^k$$

Where c is a constant term, α and φ are the coefficients of the variables that capture regional size effects, the coefficient of the interaction variables $y[j]_i z[j]^k$ is $\beta[j]$, while the coefficients of $z[j]^k$ and $y[j]_i$ are $-\beta[j]\psi[j]$ and $-\beta[j]\phi[j]$, respectively.

However, in recent literature, another version of this specification is considered, since it introduces dummies to capture unobserved characteristics. In this way, in this study we are following the estimating equation proposed by Klein and Crafts (2012) (with the difference that we use the shares of output rather than the shares of employment as the dependent variable, in the same way of Midelfart-Knarvik et al. (2000), which also has been applied by Ronsse and Rayp (2015), Nikolic (2018), and Missiaia (2019)). The specification is as follows:

$$\ln(s_i^k) = c + \sum_i \alpha_i\, Region_{i,t} + \sum_k \delta_k\, Industry_{k,t} + \sum_j \beta_j\, Interaction_{i,k,t} + \xi_i^k$$

Where s_i^k is the same dependent variable described before and c is a constant. The second and the third components are dummy variables introduced for each region and industry to control for size differences among regions and industries and any other unobserved characteristics. The variable *Interaction* is a vector of the interactions between regional characteristics and industry characteristics. β_j is the same $\beta[j]$ mentioned before and will be the focus of our analysis.

For the present study, the specific estimated equation is as follows:

$$\begin{aligned}\ln(s_i^k) = c + \sum_i \alpha_i Region_{i,t} + \sum_k \delta_k Industry_{k,t} \\ + \beta_1(Agriculture\,production \times Agricultural\,input\,use)_{i,k,t} \\ + \beta_2(Electricity\,abundance \times Electricity\,intensity)_{i,k,t} \\ + \beta_3(Educated\,population \times Skilled\,labor\,intensity)_{i,k,t} \\ + \beta_4(Credit\,abundance \times Capital\,intensity)_{i,k,t} \\ + \beta_5(Market\,potential \times Sales\,to\,industry)_{i,k,t} \\ + \beta_6(Market\,potential \times Intermediate\,goods)_{i,k,t} \\ + \beta_7(Market\,potential \times Size\,of\,establishment)_{i,k,t} + \xi_i^k\end{aligned} \quad (1)$$

The first four interactions are included considering the Heckscher–Ohlin (H–O) theory based on factor endowments, while the last three interactions consider the components described by the New Economic Geography (NEG) theory.

Following Klein and Crafts (2012), this equation can be estimated using OLS by pooling across the periods of study or by considering a single period. The estimation issues, such as heteroskedasticity or endogeneity, are discussed in Section 6.

Now we continue describing the specific variables that describe the characteristics of regions and industries and the interactions to consider for our study, taking into account the H–O theory and the NEG theory (Table 1).

The first interaction is between the agricultural production in the region, measured as a percentage of the GDP of the region, and the agricultural input as a percentage of the output of the industry.

Table 1
Regional and industry characteristics, and interactions

Framework	**Interaction**	**Regional characteristic**	**Industry characteristic**
Heckscher–Ohlin (H–O) theory	1. Agricultural factor	Agriculture production (% of GDP)	Agricultural input (% total costs of output)
	2. Electrical energy	Electricity abundance	Electricity intensity
	3. Human capital	Educated population (% of the population with secondary education)	Skilled labor intensity
	4. Financial capital	Credit abundance	Capital intensity
New Economic Geography (NEG) theory	5. Backward linkages	Market potential	Sales to industry (% of output)
	6. Forward linkages	Market potential	Intermediate goods (% of output)
	7. Economies of scale	Market potential	Size of establishment

The second interaction is between the electricity abundance in a region, measured by the inverse of the electricity price (1/*electricity price*), and the electricity intensity in the industry. The first variable informs us about the abundance of electricity in a region under the assumption that a higher price of a factor in a region represents a relatively scarce resource in relation to other regions; in contrast, a lower price represents a relatively abundant resource. The electricity intensity is measured by the average of the total equipment power per establishment in the industry. The reason to include electricity is that the Peruvian Economic Censuses of the years 1963 (Dirección Nacional de Estadísticas y Censos, 1963) and 1974 (Instituto Nacional de Estadística e Informática, 1977) reveal two main sources of industrial energy: fuels, which include firewood, coal, gas oil, gasoline, lubricants and other related products (Dirección Nacional de Estadísticas y Censos, 1963), and electricity. In 1963, electricity accounted for 21 % of industrial energy use, and 42 % in 1974.

The third interaction is between the level of educated population, which is the share of the population in the region with secondary education, and the skilled labor intensity in the industry, which is the percentage of skilled workers in relation to the total employees.

The fourth interaction is between the abundance of capital and the capital intensity. The former is measured by the share of the total amount of credit in Peru allocated to the region. Capital intensity is measured by the ratio between the value of the assets of the industry and the value of the production in the industry. It is worth mentioning that in contrast to other studies (for example, Midelfart-Knarvik et al. 2000 and Crafts and Mulatu 2005) the assumption of a high degree of capital mobility has not been considered since the accidental geography of Peru and the relatively modest development of the transport infrastructure among regions during this period avoid us to consider it as a plausible assumption.

The fifth interaction considers the market potential in regions and the sales to the industry as a percentage of the output of the industry. The variable market potential in regions is an estimation, which is explained in the following paragraphs. The sixth interaction is between the market potential and the percentage of intermediate goods that the industry uses in its production. Finally, the seventh interaction is between the market potential and the economies of scale present in the industry, the latter one measured by the size of the establishment, by considering the average number of employees per establishment in the industry.

A log transformation was applied to the variables *Agriculture production*, *Electricity abundance*, *Educated population*, *Credit abundance*, *Capital intensity*, *Market potential* and *Size of establishment* to reduce skewness, and to linearize the interaction variable with the dependent variable given that we are using OLS for modelling.

Based on the estimation of the market potential proposed by Harris Harris (1954), Crafts (2004) elaborates a methodology to estimate the market potential of a region:

$$MP_i = \sum_j \frac{GDP_j}{d_{ij}} + \sum_k \frac{foreign\_GDP_k}{d_{ki}} + \frac{GDP_i}{d_{ii}}.$$

Where MP_i is the market potential of region i. GDP_j is the GDP of the neighboring region j, divided by the distance between region i and region j (d_{ij}). $foreign\_GDP_k$ is the GDP of the main trading partners of the country, which is divided by the distance between region i and country k (d_{ki})[1]. GDP_i is the GDP of region i and d_{ii} is the internal distance within the region. To obtain the latter, the following estimation is as follows:

$$d_{ii} = 0.333\sqrt{\frac{area\ of\ region_i}{\pi}}.$$

Where $area\ of\ region_i$ is the number of squared kilometers (km^2) of region i, and the result reflects the internal distance as one third of the radius of a circle with the same area as region i.

The idea is to measure what are the accessible markets from region i, considering the economic activity of the neighboring regions, the main trading partners, and distances, in which the latter are taken as a factor of discount concerning the attractiveness of that accessibility. The third component adds the "self-potential" of the region generated by its own economic activity, discounted by an internal distance that considers how large the region is.

5. Data

The majority of data for the present investigation has been extracted from the historical documents of the Peruvian Economic Censuses of the years 1963 (Dirección Nacional de Estadísticas y Censos, 1963) and 1974 (Instituto Nacional de Estadística e Informática, 1977). Hereafter, Economic Censuses, which provide us with information of 22 industries for 23 Peruvian regions. For these years, the original data grouped industries according to the International Standard Industrial Classification of All Economic Activities (ISIC) Rev. 1, and for the year 1974, under the ISIC Rev. 2.

For the Chow-F Test, it was necessary to standardize the industry classification. For that aim, we harmonized the codes by considering the correspondence of the major groups in ISIC Rev. 1 and ISIC Rev. 2, according to the information provided by the United Nations (1958, 1968).

Concerning our dependent variable, we use the gross output of each industry per region available in the Economic Censuses. The population for the years 1963 and 1974 was calculated through a linear interpolation using the population data taken from the 1961 (Instituto Nacional de Planificación, 1964), 1972[2], and 1981 (Instituto Nacional de Estadística e Informática, 1982) Peruvian Censuses of Population. Data about the share of regional manufacturing output in total national manufacturing output is obtained from the Economic Censuses.

Regarding data on the agricultural sector, the regional agricultural production as a share of the GDP of the region is obtained from Seminario et al. (2020), by considering the production that is destined for the internal market. To estimate the agricultural input used by industries, we used the input-output tables for the years 1969 and 1973 published by the Instituto de Planificación/Instituto Nacional de Estadística, which are reproduced in the work of Gonzales de

[1]Following Nikolic (2018), to measure the distance between regions or foreign countries and regions, it is considered in general the distance between the capital of the region or the country, since they are the center of market activities.

[2]Even though the 1972 census is not published online, it was possible to obtain the number of the population per region because the 1981 population census also reports that information (Instituto Nacional de Estadística e Informática, 1982).

Olarte (2016). It is worth mentioning that we are following the assumptions of other authors in which the intensity in the use of factors is representative for closer periods (for example, (Martínez-Galarraga, 2012) and (Ronsse and Rayp, 2015)), and similarly, we consider the 1969 input-output table to be representative of the year 1963 while the table for 1973 is considered the reference for 1974.

The price of electricity is obtained from the Economic Censuses by considering the ratio of the reported amount of total purchased energy in kilowatt-hour (kWh) in the region and the total amount paid for this purchase in local currency (soles) by establishments with more than five people. The value of the total equipment power in horsepower units in the industry per region is also obtained from the Economic Censuses, which is then converted into kWh using the Peruvian conversion factor of 0.7355.

The number of individuals with secondary school per region is obtained from a linear interpolation by considering the population with secondary school in the years 1961 and 1981, according to the Peruvian Censuses of Population of each year (Instituto Nacional de Planificación (1964) and Instituto Nacional de Estadística e Informática (1982), respectively). The information about the total number of skilled employees in each industry is obtained from the Economic Censuses.

The amount of credit per region is obtained from the reports of the Superintendence of Banking of Peru (Superintendencia de Bancos, 1965; Superintendencia de Banca y Seguros, 1983). The credits of private banks and the state-owned Industrial Bank were considered since the latter was created to promote industrial activity (Banco Central de Reserva del Perú, 1936). Even though the data available for 1974 was not published, it was possible to generate an interpolation of the values from the reports mentioned. Although there is no Consumer Price Index published by the government for the period of the analysis, we used the Gross Fixed Capital Formation Price Index estimated by Seminario (2016) to convert the credit values into constant 1979 prices, under the assumption that credits were directed to invest in assets. The number of establishments per industry in each region is obtained from the Economic Censuses. The values of assets (machinery and equipment) for each industry in each region are obtained from the Economic Censuses.

To calculate the market potential, we use the regional GDP expressed in 1990 international Geary–Khamis dollars estimated by Seminario et al. (2020) for Peruvian regions, while the distances between neighboring regions are obtained from the Instituto Nacional de Estadística e Informática (1996b). For the GDP of the trading partners, we consider the five most important partners of Peru, which generated 75 % of the trade value for the years 1966 and 1969 (Ministerio de Hacienda y Comercio, 1969; Dirección Nacional de Estadísticas y Censos, 1963). These are the United States, Japan, West Germany, Argentina and the Netherlands. The GDP for each country is obtained from Maddison (2003) and is expressed in 1990 international Geary–Khamis dollars. To measure the distance (km), we use the sea distance obtained from https://sea-distances.org/ and to measure the locations connected by land, we use https://www.google.com/maps (for example, when a region or a city was landlocked, we calculate the distance to the closest port).

The percentage of sales to industries is obtained from the input–output tables of 1969 and 1974 mentioned before. The percentage of intermediate goods that industries use in their production is obtained from the Economic Censuses. To measure the size of the industry, the number of employees and the number of establishments per region are obtained from the Economic Censuses.

Among the limitations, we found that, for the year 1963, the data for industries were divided into two groups based on the number of employees per establishment, with a threshold of 5 people. For the establishments with more than 5 people, the data was disaggregated in detail, but for establishments with fewer than 5 people, the data was aggregated. Even though the data available for establishments with fewer than 5 people makes it difficult to generate a reliable estimation for each variable of the model, it is considered that not using this data does not affect our estimation, since their value of production for 1963 constituted 2 % of the total value of production in the Peruvian manufacturing sector.

Another limitation to consider is that the fuels, such as firewood, coal, gas, oil, gasoline, lubricants and other related products were not considered in our model due to the unavailability of disaggregated data for these inputs. However, instead, we only use electricity under the current circumstances and we consider it to be the best alternative to measure industrial energy sources since it represents a relevant input in the production function of industries, as mentioned before.

Finally, some statistics are presented to provide a broader context about the level of industrialization among Peruvian regions for the period of study. In Table 2, for 1963, from a total of 25 321 industrial establishments, approximately 33,9 % of them were placed in Lima, 10,2 % in Piura, and 6,5 %. For 1974, the total number of industrial establishments amounted to 25 156, of which 36,4 % were placed in Lima, 8,2 % in Junin and 6,1 %. What is remarkable is the persistent presence of the majority of industrial establishments in the Lima region in 1963 and 1974.

Table 2

Share of manufacturing industrial establishments in Peruvian regions

1963		1974	
Lima	33.9 %	Lima	36.4 %
Piura	10.2 %	Junín	8.2 %
Junín	6.5 %	Cusco	6.1 %
Cajamarca	5.8 %	Cajamarca	5.7 %
Arequipa	5.8 %	Piura	5.5 %
Cusco	5.5 %	Ancash	5.5 %
Ancash	4.5 %	La Libertad	5.0 %
La Libertad	4.4 %	Arequipa	4.9 %
Lambayeque	4.4 %	Lambayeque	4.3 %
Ica	2.9 %	Puno	2.6 %
Ayacucho	2.7 %	Ica	2.4 %
Puno	2.5 %	Loreto	1.8 %
Huánuco	2.0 %	Huánuco	1.8 %
Loreto	1.9 %	San Martín	1.5 %
Amazonas	1.5 %	Amazonas	1.5 %
Apurímac	1.1 %	Ayacucho	1.3 %
Tacna	0.9 %	Tacna	1.2 %
Pasco	0.9 %	Pasco	0.9 %
Huancavelica	0.8 %	Moquegua	0.7 %
San Martín	0.8 %	Apurímac	0.7 %
Tumbes	0.7 %	Huancavelica	0.6 %
Moquegua	0.5 %	Tumbes	0.6 %
M. de Dios	0.1 %	M. de Dios	0.1 %

Source: Dirección Nacional de Estadísticas y Censos (1963) and Instituto Nacional de Estadística e Informática (1977)

Table 3 shows the indicators of the regional characteristics considering the average of the values for 1963 and 1974. It includes the share of the regional GDP that represents the agricultural activity, the electricity price, the share of the population in the region with secondary school education, the share of credits allocated in Peruvian regions, and the market potential estimated per region.

In Table 4, we present the indicators of industry characteristics considering an average of the values for the years 1963 and 1974. Here, we present the extent to which the agricultural inputs, asset values, the sales to other industries, and the intermediate goods used in the industry represent each industry's total value of production. Additionally, the table includes information about the average power per plant, the percentage of skilled labor and the average number of employees per industry.

Table 3
Regional characteristics (Average of 1963 and 1974)

Region	Agriculture production (% of regional GDP)	Electricity price (Sol per kWh)	Secondary school education (% of population)	Credit abundance (% of total credit in Peru)	Market potential*
Amazonas	20 %	0.72	5 %	0.02 %	765
Ancash	10 %	4.19	8 %	2.69 %	803
Apurímac	18 %	0.78	4 %	0.05 %	662
Arequipa	4 %	1.75	17 %	3.72 %	711
Ayacucho	15 %	1.16	5 %	0.14 %	685
Cajamarca	24 %	0.98	4 %	0.33 %	771
Cusco	14 %	1.69	7 %	1.09 %	572
Huancavelica	14 %	1.00	4 %	0.05 %	782
Huánuco	18 %	0.89	6 %	0.36 %	757
Ica	5 %	1.43	16 %	1.99 %	749
Junín	7 %	1.09	11 %	1.30 %	749
La Libertad	10 %	0.90	11 %	2.37 %	760
Lambayeque	7 %	1.33	13 %	1.36 %	797
Lima	1 %	1.58	23 %	79.94 %	1467
Loreto	7 %	1.59	6 %	1.48 %	607
Madre de Dios	7 %	—	17 %	0.02 %	617
Moquegua	7 %	0.98	17 %	0.10 %	654
Pasco	5 %	0.87	8 %	0.03 %	772
Piura	6 %	1.41	8 %	2.03 %	851
Puno	20 %	0.85	5 %	0.38 %	593
San Martín	10 %	0.90	8 %	0.11 %	676
Tacna	3 %	1.05	20 %	0.36 %	620
Tumbes	3 %	1.41	14 %	0.07 %	814

* Millions of 1990 international Geary–Khamis dollars.
Source: See Section 4.

Table 4

Industry characteristics (Average of 1963 and 1974)

Industry	Agricultural input (% of value)	Electricity intensity (kWh per plant)	Skilled labor intensity (% employees)	Capital intensity (% of value)	Sales to industry (% of value)	Intermediate goods used (% of value)	Size of establishment (avg. employees)
Food manufacturing industries	26 %	69	12 %	18 %	19 %	54 %	18
Beverage industries	5 %	221	20 %	22 %	6 %	35 %	16
Manufacturing of textiles	14 %	71	9 %	28 %	39 %	43 %	21
Manufacture of footwear	0.2 %	4	14 %	13 %	0.3 %	49 %	6
Manufactures of wood and cork	5 %	237	11 %	34 %	19 %	45 %	10
Manufacture of furniture	5 %	12	10 %	16 %	19 %	45 %	5
Manufacture of paper	0 %	1124	27 %	75 %	89 %	46 %	112
Printing and publishing industries	0 %	36	31 %	43 %	6 %	38 %	10
Manufacture of leather and fur products	32 %	194	21 %	24 %	74 %	54 %	11
Manufacture of rubber products	3 %	1477	29 %	15 %	27 %	49 %	11
Manufacture of chemical products	0.4 %	290	31 %	49 %	53 %	45 %	27
Manufacturing of non-metallic products	0.03 %	99	12 %	35 %	23 %	33 %	13
Basic metal industries	0.02 %	5730	21 %	33 %	89 %	49 %	791
Manufacture of metal products	0.01 %	514	19 %	31 %	35 %	47 %	9
Manufacture of machinery	0.01 %	23	25 %	80 %	54 %	40 %	15
Manufacture of electrical machinery	0 %	671	35 %	24 %	25 %	46 %	21
Manufacture of transport equipment	0.05 %	28	16 %	32 %	27 %	49 %	13

Note: For this table, two outlier observations were removed regarding capital intensity: one for the manufacturing of non-metallic products and another for the basic metal industries.

Source: See section 4.

6. Estimation and results

In this section, we present the results of our estimations. First, we determine if it is possible to pool the data of the two periods by performing a Chow–F Test, which allows us to identify if the coefficients remain stable for both years (Gujarati and Porter, 2009). The result of the test (F(44,371) = 5.2827) indicates that the null hypothesis of stable parameters is rejected at the 1 % significance level and therefore suggests that the estimation should be generated per year and not pooled. For this period, it is plausible to consider that this result responds to the change Peru passed through since a dictatorship was established by the Revolutionary Government of the Armed Forces of Peru in 1968, breaking with the democracy Peru was living in at that time. The rise of this new regime came with new political measures trying to generate a deep impact and change in the social and economic bases of the country.

As a next step, it is important to consider two potential issues associated with the type of data used for this study, which are heteroskedasticity and endogeneity.

There is a potential problem of heteroskedasticity since our data contains different industries and regions. Additionally, since more than one type of industry is also present in different regions of Peru, it is possible that our results can be subject to an unobserved cluster effect due to regional differences within the country. Overlooking these factors could also impact our estimations (Pepper, 2002). For this reason, to address these issues, we present our results using an OLS estimation with White's heteroskedasticity-consistent standard errors and regional cluster-robust standard errors.

It is worth mentioning that, following Midelfart-Knarvik et al. (2000, 2001), the estimations exclude the industry of petroleum refineries (given that their locations are based mainly on the locations of this natural resource) and manufacturing classified as "other manufacturing industries" for being a residual component. Regarding the former, we verify, for example, that during the period of analysis, Piura, the main oil-producing region, absorbs on average 42 % of the refinery production of the country. The estimation results are presented in Table 5.

For both years 1963 and 1974, the regression models are well-fitted, with R^2 values of 0.8837 and 0.8647, respectively. Region and industry dummies are included in every regression to control for size differences among regions and sectors, as well as for unobserved characteristics.

For the year 1963, the coefficients of the first, second, fourth, fifth and seventh interactions are significant. For interaction 1 *(Agriculture production × Agricultural input use)*, the coefficient is significant at the 10 % level when using robust standard errors and cluster-robust standard errors. This means that regions where agricultural production as a share of the regional GDP was high resulted in attractive locations for industries that used a high level of agricultural inputs.

The interaction 2 *(Electricity abundance × Electricity intensity)* is statistically significant at the 1 % level when using robust standard errors and cluster-robust standard errors. It means that in 1963, industries with intense use of electricity in their establishments preferred regions where electricity was abundant.

The interaction 4 *(Credit abundance × Capital intensity)* is significant at the 1 % significance level when using robust standard errors and cluster-robust standard errors, which means that in

Table 5
Estimation results (Dependent variable: $\ln(s_i^{k*})$)

	1963		1974	
Variables	(i)	(ii)	(i)	(ii)
Interaction 1				
Agric. Production × Agric. input	1.156*	1.156*	1.268*	1.268*
	(0.624)	(0.662)	(0.732)	(0.742)
Interaction 2				
Elec. availab. × Elec. intensity	0.430***	0.430***	0.033	0.033
	(0.075)	(0.103)	(0.042)	(0.041)
Interaction 3				
Educ. pop. × Sk. lab. intensity	-0.270	-0.270	-0.110	-0.110
	(0.322)	(0.203)	(0.042)	(0.436)
Interaction 4				
Cred. availab. × Cap. intensity	0.063***	0.063***	-0.001	-0.001
	(0.015)	(0.017)	(0.018)	(0.019)
Interaction 5				
Market potential × Sales to ind.	1.751**	1.751**	1.646**	1.646**
	(0.776)	(0.621)	(0.835)	(0.566)
Interaction 6				
Market potential × Inter. goods	0.043	0.043	-0.025	-0.025
	(0.039)	(0.031)	(0.031)	(0.023)
Interaction 7				
Market potential × Size of est.	0.045***	0.045***	0.047***	0.047***
	(0.007)	(0.006)	(0.006)	(0.007)
Constant	-42.764***	-42.764***	-41.110***	-41.110***
	(14.401)	(11.856)	(15.613)	(10.683)
Regional dummy	Yes	Yes	Yes	Yes
Industry dummy	Yes	Yes	Yes	Yes
Number of observations	158	158	213	213
R^2	0.8837	0.8837	0.8647	0.8647

Note: (i) Regression with robust standard errors; (ii) Regression with regional cluster-robust standard errors. Standard deviation in parentheses. * Significant at 10 %; ** significant at 5 %; *** significant at 1 %.

1963, industries with a highly intense use of capital in their production function preferred to be located in regions where credits were abundant compared to other regions.

For the case of interaction 5 *(Market potential × Sales to industry)*, the coefficient is significant at the 5 % level when using robust standard errors and cluster-robust standard errors, meaning that in 1963, industries with high levels of sales among industries consider regions with high market potential to be attractive places to locate.

Regarding interaction 7 *(Market potential × Size of establishment)*, the coefficient is significant at the 5 % level when using robust standard errors and at the 1 % level when using cluster-robust standard errors. It implies that in 1963, industries that present economies of scale found attractive places with high market potential.

For the year 1974, the coefficients of the first, fifth and seventh interactions are significant. Regarding interaction 1, in this case, the coefficient is statistically significant at the 10 % significance level when using robust standard errors and cluster-robust standard errors. Considering the significant result, it denotes the same relation for the year 1963 between agricultural production in the region and the intensity of agricultural inputs used in the industry.

For interaction 5, similar to 1963, the coefficient is also significant at the 5 % level when using robust standard errors and at the 1 % level when using cluster-robust standard errors, implying that the same relationship described for 1963 remains in this year.

Regarding interaction 7, the coefficient is significant at the 1 % level when using robust standard errors and cluster-robust standard errors, which implies that the same relationship described for 1963 also remains in this year.

Even though these results, it is important to consider that the variable market potential can generate a problem of endogeneity in our model. The problem arises because a location where the level of economic activity is high will attract more industries and, as a result, will increase the market potential of the region (Martínez-Galarraga, 2012). A problem like this in the specification implies that the independent variable can be correlated with the unobservable random error and therefore can generate biased and inconsistent estimators (Cameron and Trivedi, 2005; Gujarati and Porter, 2009; Wooldridge, 2018).

To address this issue, we estimate our model again using an instrumental variable with a two-step General Method of Moments (GMM) estimation. This method allows us in the first stage to remove the correlation of the endogenous explanatory variable and the unobservable random error (Wooldridge, 2020), which is then used in a second-stage regression to obtain consistent estimators.

For our instrumental variable, we follow the empirical strategy of Klein and Crafts (2012), who consider as instrumental variable the distance to the eastern seaport, New York City. In our case, the instrumental variable is the distance of regions to the main economic center of Peru, Lima, according to its Regional GDP (see Figure 1), which also contains the main port of the country. In that way, being closer to this city represents high access to markets.

It is worth noting that Klein and Crafts (2012) mention that New York City, being an economic center long before the period of their study, might have influenced the geographical distribution of manufacturing industries across the United States. A similar situation can occur with Lima for the case of Peru. However, given the limitations imposed by the natural accidental geography of the country and the limited connectivity among regions, it is plausible to consider that industries could have preferred to be placed in Lima instead of nearby regions. Additionally, the expansion of the process of industrialization in Peru is not long before our period of study, since, according to Vega-Centeno (1988), it started in the 1950s. For this reason, we consider that this variable fulfills the conditions to be an instrument of market potential.

Table 6
GMM first-stage instrumental variable estimation results
(Dependent variables: Interactions 5, 6 and 7)

	1963			1974		
Variables	Int. 5	Int. 6	Int. 7	Int. 5	Int. 6	Int. 7
Interaction 1: Agric. production × Agric. input						
	−0.124	−1.736***	2.828	−0.038	−1.728**	6.149
	(0.027)	(0.473)	(2.232)	(0.023)	(0.715)	(4.423)
Interaction 2: Elec. availab. × Elec. intensity						
	−0.002	−0.011	0.066	−0.0003	−0.035**	0.063
	(0.002)	(0.023)	(0.152)	(0.001)	(0.018)	(0.138)
Interaction 3: Educ. pop. × Sk. lab. intensity						
	0.039***	−0.118	0.463	0.006	−0.048	−1.166
	(0.012)	(0.098)	(0.710)	(0.010)	(0.192)	(1.224)
Interaction 4: Cred. availab. × Capital intensity						
	−0.0004	0.008***	−0.014	0.0002	−0.012	0.042
	(0.0004)	(0.004)	(0.044)	(0.0007)	(0.007)	(0.051)
Interaction 5 (IV): Distance to Lima × Sales to indus.						
	−0.201***	0.085	−1.017	−0.240***	−0.618**	−3.442**
	(0.010)	(0.170)	(0.945)	(0.010)	(0.269)	(1.701)
Interaction 6 (IV): Distance to Lima × Interm. goods						
	−0.0006	3.060***	0.044	0.003	3.141***	−0.111
	(0.003)	(0.043)	(0.192)	(0.002)	(0.041)	(0.244)
Interaction 7 (IV): Distance to Lima × Size						
	−0.0005	0.013*	3.134***	−0.0004	0.016**	3.147***
	(0.008)	(0.007)	(0.093)	(0.0004)	(0.008)	(0.062)
Number of observations	158	158	158	213	213	213
F-statistic on excluded instruments	148.72	1742.31	441.51	191.23	1985.67	1007.05
Sanderson–Windmeijer F-statistic	458.36	4374.42	1586.26	654.76	5968.79	2674.84

Note: Regression with robust standard errors in parentheses. * Significant at 10 %; ** significant at 5 %; *** significant at 1 %.

We apply the test proposed by Stock and Yogo (2005) to evaluate whether this variable is a weak instrument (which can reduce the consistency and efficiency of the OLS estimators). To apply the test, we calculate the Cragg–Donald F-statistic and compare it with the critical values defined by Stock and Yogo. The results of the F-statistic are displayed in Table 7 and indicate that for every year, it is possible to reject the hypothesis that the instrumental variable is weak.

Regarding the estimation, we use a two-step GMM instrumental variable estimation with robust standard errors[3], since this estimator is more efficient if heteroskedasticity is present (Baum et al., 2003). Additionally, Table 6 displays the results of the first-stage estimation.

The results are displayed in Table 7 and confirm the relevance of the first, second, fourth, fifth and seventh interactions. Regarding 1963, the coefficients of interactions 2, 4 and 7 are significant at the 1 % level, while for interaction 1 and 5, the coefficients are significant at the 5 % level. Regarding 1974, the coefficient of interaction 1 is significant at the 10 % level, and interaction 7 is significant at the 1 % level. The main difference with the estimation of Table 5 is that the interaction 5 is not significant for the year 1974.

Moreover, we display in Table 8 the standardized coefficients of the relevant interaction variables to provide an understanding of the relative weights that each interaction had over the location decision of the industries. The interactions considered are those that are significant at least at a minimum of 10 %, considering our estimation in Table 8 when using the instrumental variable.

Table 8 shows that the New Economic Geography forces exert a more important influence on industrial location decisions for the years 1963 and 1974. Regarding 1963, the interaction between the market potential with the backward linkages and economies of scale in industries has a relative impact of 86 %. Even though in 1974 the backward linkages interaction is not present, the influence of the New Economic Geography forces remains relevant, since the interaction between the market potential and the economies of scale in industries, the relative impact on industrial location decisions represents 79 %.

[3]The industry dummies are partialled out from both the potentially endogenous regressors and the instruments to ensure the full rank of the covariance matrix of moment conditions.

Table 7
GMM Second-stage instrumental variable estimation results
(Dependent variable: $\ln(s^*_{ik})$)

	Year	
Variables	1963	1974
Interaction 1		
Agric. Production × Agric. input	1.156**	1.189*
	(0.533)	(0.645)
Interaction 2		
Elec. availab. × Elec. intensity	0.434***	0.037
	(0.064)	(0.038)
Interaction 3		
Educ. pop. × Sk. lab. intensity	-0.270	-0.116
	(0.277)	(0.392)
Interaction 4		
Cred. availab. × Capital intensity	0.0626***	0.0002
	(0.013)	(0.016)
Interaction 5 (IV)		
Distance to Lima × Sales to indus.	1.676*	1.089
	(0.779)	(0.793)
Interaction 6 (IV)		
Distance to Lima × Interm. goods	0.038	-0.024
	(0.034)	(0.025)
Interaction 7 (IV)		
Distance to Lima × Size	0.044***	0.045***
	(0.006)	(0.006)
Number of observations	158	213
Regional dummy	Yes	Yes
Industry dummy	Yes	Yes
Cragg–Donald F-statistic	197.461	306.858
R^2	0.8646	0.8447

Note: Regression with robust standard errors (in parentheses). * Significant at 10 %; ** significant at 5 %; *** significant at 1 %.

Table 8
Standardized beta coefficients

Framework	**Interaction**	Standardized beta coefficients	
		1963	**1974**
Heckscher–Ohlin (H–O) theory	*Interaction 1* Agricultural factor	0.148** (3 %)	0.152* (21 %)
	Interaction 2 Electrical energy	0.319*** (7 %)	
	Interaction 4 Financial capital	0.219*** (4 %)	
New Economic Geography (NEG) theory	*Interaction 5* Backward linkages	3.778** (77 %)	
	Interaction 7 Economies of scale	0.424*** (9 %)	0.569*** (79 %)

Note: *Significant at 10 %; **significant at 5 %; ***significant at 1 %. Relative share in parentheses.

In particular, our findings highlight that the factors that influence industrial location decisions in Peruvian regions are a mix of the factors that Heckscher–Ohlin and New Economic Geography theories predict. This is aligned with studies that find the simultaneous influence of both factors in industrial location decisions, such as Crafts and Mulatu (2005), Wolf (2007), Sanguinetti and Volpe Martincus (2009), Klein and Crafts (2012), Badia-Miró and Yáñez (2015), Nikolic (2018), and Missiaia (2019).

The relevant interactions in our study are also relevant in the case of other countries. For the case of interaction 1, regarding the agricultural factor, the results are aligned with the findings of Crafts and Mulatu (2005) and Martínez-Galarraga (2012). For interaction 2, regarding electrical energy, it is also relevant for the cases of the United Kingdom (Crafts and Mulatu, 2005), Spain (Martínez-Galarraga, 2012) and Italy (Missiaia, 2019). Interaction 5, regarding backward linkages, has also been highlighted as relevant for the case of the United States (Klein and Crafts, 2012).

In relation to interaction 6, regarding economies of scale, the results are aligned with the findings of **?**, Martínez-Galarraga (2012), Klein and Crafts (2012), Badia-Miró and Yáñez (2015), and Ronsse and Rayp (2015). The only exception is for interaction 4, the interaction regarding financial capital, since no case identifies the interaction of financial capital as significant; however, it is important to mention that only Nikolic (2018) and Missiaia (2019) consider this factor in their models.

7. Discussion

Now, regarding the particularity of the Peruvian case, we can provide the following interpretation of the findings. First, it is worth noting that for both years, the agricultural abundance in regions attracts industries intensive in agricultural inputs, and the market potential of regions attracts industries with economies of scale.

The particularity here is that the significance of the agricultural interaction changes from 5 % in 1963 to 10 % in 1974 (see Table 8), while the interaction related to the economies of scale remains significant at the 1 % level for both years. This pattern could be similar to the one faced in Chile (Badia-Miró and Yáñez, 2015), in which the relevance of the endowment factors was relevant before 1930, but after that, industries with economies of scale clustered in provinces with high market potential.

For the Peruvian case, this can respond to a progressive reduction in transportation costs to mobilize agricultural inputs from rich-source regions, allowing firms to be placed near markets with high potential. According to Matos Mar (2010), around 1950, paved and gravel roads substituted the former system of bridle paths and footpaths, generating a new road network that increased the connectivity of Peruvian regions, which can explain the reduction in transportation costs.

One concern for this argument is that the railway network already existed since the second half of the nineteenth century in Peru, and therefore transportation costs could already have been low before the 1950s. This is precisely what Martínez-Galarraga (2012) finds in his study for Spain between 1856 and 1929, explaining that the railway network was a factor that influenced on the spatial concentration of industries. However, the Peruvian rail system presents a particularity: it was not originally designed to connect rich-source regions and regions with attractive market potential; it was designed to extract natural resources from the country to export them abroad. According to Ministerio de Transportes y Comunicaciones (2021), before the mid-twentieth century, the Peruvian economy presented an isolated and specific transport network whose main function was to connect production centres (mines, country estates, or agricultural spots) with exportation hubs, and neglecting in this way the development of the internal market. For example, in the beginning of the twentieth century, it is reported that trail cars mainly transported minerals, wool and sugar and other products for exportation (Ministerio de Transportes y Comunicaciones, 2021). This is closely related to the extractivist economic model established since the Spanish Colony, with descriptions of this pattern persisting even after the country achieved its independence in 1821.

Another evidence of this is the contract signed by the Peruvian Government and Enrique Meiggs to build one of the largest railway lines in Peru, the Central Andean Railway. Part of the supportive arguments cited in the contract is the study "Estudio sobre la provincia de Jauja" (Pardo y Lavalle, 1862), which elaborates the idea of building a train that connects the Junín region rich in minerals with exportation hubs on the Peruvian coast.

The statistics about the total amount of weight mobilized in the railway system also provide information about the role of this network in the Peruvian economy. According to the report provided by the Oficina Nacional de Estadística y Censos (1975), the most important product mobilized in the Peruvian railway system from 1960 to 1969 was minerals, representing on average 52 % of the total weight mobilized, while livestock and animal products accounted for 2 %, agricultural products 10 %, forest products 2 %, fuels 11 %, and other products 22 % (see Tables 9 and 10). This information reveals the dominant role of mineral transportation in the railway system, supporting the idea that this network did not focus on reducing the transport costs of industries, but instead on extracting natural resources for exportation.

Table 9

Total amount of weight mobilized in the railway system by category of products 1960–1969 (In thousands of metric tons)

Category	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
Livestock and animal products	112	115	101	99	107	91	76	48	42	41
Agricultural products	584	542	476	476	489	502	457	231	159	170
Forest products	116	121	103	98	111	101	106	67	49	46
Minerals	1927	1917	1864	2037	2192	2314	2349	2127	1874	1729
Fuels	405	411	403	401	424	427	442	461	435	438
Others	895	884	806	760	884	890	842	921	808	773
TOTAL	**4039**	**3990**	**3753**	**3871**	**4207**	**4325**	**4272**	**3855**	**3367**	**3197**

Source: Oficina Nacional de Estadística y Censos (1975, p. 822).

Table 10

Percentage of the total weight mobilized in the railway system by category of products 1960–1969 (%)

Category	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
Livestock and animal products	2,8	2,9	2,7	2,6	2,6	2,1	1,8	1,2	1,3	1,3
Agricultural products	14,5	13,6	12,7	12,3	11,6	11,6	10,8	6,0	4,7	5,3
Forest products	2,9	3,0	2,7	2,5	2,6	2,4	2,5	1,7	1,5	1,4
Minerals	47,7	48,0	49,7	52,6	52,1	53,5	55,0	55,2	55,6	54,1
Fuels	10,0	10,3	10,7	10,4	10,1	9,9	10,3	12,0	12,9	13,7
Others	22,1	22,2	21,5	19,6	21,0	20,5	19,6	23,9	24,0	24,2
TOTAL	100	100	100	100	100	100	100	100	100	100

Source: Oficina Nacional de Estadística y Censos (1975, p. 823)

For this reason, we consider that the railway system, in contrast to the Spanish railway system, did not generate a reduction of the transport cost that industries faced during our period of study.

Regarding the influence of electrical energy in the industrial location, it is noticeable that this factor turned into a non-significant variable in 1974. This can be the result of the nationalist military regime that seized control of the country in 1968, which, among its policies, nationalized the energy sector to promote a decentralized national industrialization. Given that in 1972 the government's share in electric power plants changed, rising from 9.7 % to 98.1 % (Osinergmin, 2016), it is possible to consider that the orientation of the government generated a decentralized distribution of energy, reducing the relevance of this factor in the industrial location decisions.

The change in the relevance of financial capital for 1974 can also be attributed to the policies implemented by the military regime. According to Lobo Collantes (2022), the state-owned Industrial Bank that provided more than 50 % of the credit for the industrial sector in the 1970s lost importance in its role of promoting the industrial sector due to the emergence of a state-owned bank for the agricultural sector, the Agrarian Bank, which assumed a greater role in the military regime's policy. This possibly brings as a result that credit lacks relevance in industrial location decisions.

The change in the relevance of backward linkages can be explained by the distortion generated by the tariffs imposed to protect local industries during the 1970s. During the 1960s, the manufacturing industry based its dynamism on the production of durable goods, which was dependent on imported capital and imported intermediate goods (Thorp and Bertram, 2020; Jiménez et al., 1999). However, to implement the process of Import Substitution Industrialization (ISI), in its aim to promote national production, the Peruvian government increased the tariffs on imported consumer products, imported intermediate goods, and imported capital equipment, which affected industrialization, since local producers were not capable of providing those types of goods (Jiménez et al., 1999). This generated a distortion in the market and probably in the supply chain, meaning that industries reduced their demand for nationally produced intermediate products due to the lack of other intermediate goods (imported ones) to elaborate their final products. As a result, this reduces the power of backward linkages to explain industrial location decisions.

Regarding the influence of economies of scale in the location patterns of manufacturing, it is worth noting that our findings align with the information provided by official sources about the development of manufacturing activity in the country. For example, according to the Central Reserve Bank of Peru, the production of textiles grew steadily for fifteen years since 1945 due to the population's increasing familiarity with mass-produced clothing and the more affordable prices that this process generated. Additionally, the Central Reserve Bank of Peru also reported the rise of the printing and publishing industries and the trend of gathering big volumes of production by big mass manufacturing workshops starting to soar during 1959 (Banco Central de Reserva del Perú, 1961).

Another relevant insight is that the factor of human capital does not have a relevant influence on the industrial location decisions. For both years, 1963 and 1974, the coefficient of this interaction is not significant, which may indicate that the industry locations were not necessarily defined by the amount of skilled labor in regions. This result can be associated with the fact that the share of skilled workers in the majority of manufacturing industries, during the years 1963 and 1974, represented one-third or less of the total workforce (see Table 4).

Additionally, the findings of this study and historical events in Peru seem to point that the agglomeration of industries in regions, but in particular in Lima, responds to a spontaneous order rather than to predetermined political decisions. Following Matos Mar (2010), the arrival of the radio in the country spread information about the relatively high standards of life in the capital Lima, and the new road system established in the 1950s provided the conditions for a massive migration from rural areas to cities. Even though internal migration to the cities was a noticeable trend during the 1920s and 1930s, this process accelerated and increased sharply during the decade of the 1950s and 1960s (Bourricaud, 2017). As a result, in 1940 the urban population rate was 35 % while the rural population rate was 65 %, but in 1960 the urban rate represented 47 % of the total population, and in 1981, the urban population rate was 65 % (United Nations, 1969; Jaramillo and Huamán, 2020; Martinelli and Vega, 2023).

Given the gradual reduction of transport costs as a result of the road system, firms were no more tied to land and preferred to be placed near areas with high market potential. This generated an agglomeration of industries near these places, generating a positive feedback process that increased the attractiveness of the location and its economic potential. Since Lima received the largest number of people during the internal migration process (Instituto Nacional de Estadística e Informática, 1995), it is plausible to consider that this factor increased its attractiveness as a place with high market potential, and probably this, in part, explains why one third of the industries agglomerated there during the period of study (see Table 2).

8. Conclusion and policy implications

In the present study, we investigate the determinants of the industrial location in Peru for the years 1963 and 1974, by using the first two Economic Censuses that the Peruvian government conducted for both years. For that aim, we apply a methodology based on Midelfart-Knarvik et al. (2000, 2001) that allows us to assess how the interaction of the regional and industry characteristics influences on the location decisions of manufacturing industries. The influence of these characteristics on the industrial location is based on the Heckscher–Ohlin (H–O) and New Economic Geography (NEG) theories.

Our findings indicate that both forces influenced simultaneously on industrial location decisions in Peru's manufacturing sector for the years 1963 and 1974. The relevant factors are related to the regional endowments of agriculture, electrical energy, and financial capital, while the significant NEG force is the interaction of the economies of scale in industries and the market potential of regions.

These factors are also relevant in other studies, except for the financial capital component. Even though financial capital has been considered before in two studies, it is the first time in the literature that the result is significant. Additionally, the contribution of this paper is that it adds information about what factors can explain the industrial location in developing countries, involving the Latin American zone.

Moreover, the findings indicate that during the period of study, the process of agglomeration of industries in some Peruvian regions, in particular in Lima, can respond to a spontaneous order rather than to predetermined political decisions, in which factors like the arrival of the radio to the country, the development of a road system, and the internal migration process, and the agglomeration of industries near regions with high market potential turned to be key. Additionally, the results also indicate that the factor of educated people was not significantly influential in the industrial location decisions, which could be explained by the share of skilled workers in industries during that period (one-third or less of the total workforce).

What is interesting is that the railway system did not seem to play a role in the reduction of transportation costs for industries, which provides an insight for further policy implications. According to the New Economic Geography theory, the level of transportation costs can generate three scenarios: a first scenario where transportation costs are high and makes industries be tied to land and a dispersion of industries occurs; a second one, in which the transportation costs are lower, allowing firms to agglomerate near areas with high market potential and no longer tied to land; and a third one, in which the transportation cost to mobilize the final products or the intermediate inputs are no longer a burden and the location is subject to labor costs, in which a dispersion of industries can emerge. Our results seem to point that for our period of study, Peru was in the process of arriving at the second scenario.

If the country were facing the third scenario, that might mean that industries would spread around regions. However, according to the official information, during the following decades after 1970s, the agglomeration of the manufacturing industry remained concentrated in Lima: From the 1990s until around the 2010s, the share of manufacturing establishments in Lima was approximately 45 % and in the rest of the regions it was less than 6 % (Instituto Nacional de Estadística e Informática, 1996a, 2010); from the 2010s until the mid-2020s, the share in Lima increased to 51 % while the share in the rest of the regions remained below 6 % (Instituto Nacional de Estadística e Informática, 2016, 2025). It seems to imply that the country remains in the second scenario described by the New Economic Geography theory, in which the level of transportation costs generates the agglomeration of industries; however, further research is required to confirm this.

If that is the case, it is relevant to note that the findings of this study indicate that the railway system in the past did not contribute to reducing the transportation cost faced by industries, as its role was to connect extraction sites with export hubs. This role appears to have remained largely unchanged even until recent years (Ministerio de Transportes y Comunicaciones, 2021; Organismo Supervisor de la Inversión en Infraestructura de Transporte de Uso Público, 2024). In this context, reassessing the role of the railway system appears as a potential option to explore for reducing the transportation costs faced by manufacturing industries, which, according to the predictions of New Economic Geography theory, might contribute to dispersing industrial activity across the Peruvian regions.

In conclusion, the present findings not only add information about the determinants of industrial location in a developing country in Latin America, but also allow us to understand the internal economic dynamics that Peru had regarding the location of the manufacturing sector in its regions for the years 1963 and 1974, and allow us to reflect about what alternatives are available to generate more even development among regions by generating the spread of the manufacturing industry.

References

Abusada, R. (1977). Políticas de Industrialización en el Perú, 1970–1976. *Economía*, 1(1):9–34.

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4):1231–1294.

Acemoglu, D. and Robinson, J. A. (2012). *Por qué fracasan los países: Los orígenes del poder, la prosperidad y la pobreza*. Planeta.

Antošová, G. and Arias Gómez, H. Y. (2024). Spatial behavior of manufacturing in Colombia: territorial case of connection ideas of Marshallian and New Economic Geography insights about industrial localisation. *Economics and Environment*, 91(4):838–838.

Arthur, W. B. (1990). Positive feedbacks in the economy. *Scientific American*, 262(2):92–99.

Badia-Miró, M. and Yáñez, C. (2015). Localisation of Industry in Chile, 1895–1967: Mining Cycles and State Policy. *Australian Economic History Review*, 55(3):256–276.

Banco Central de Reserva del Perú (1936). *Memoria 1936*.

Banco Central de Reserva del Perú (1961). *Actividades productivas del Perú 1961: Análisis y perspectivas*. Banco Central de Reserva del Perú.

Baum, C. F., Schaffer, M. E., and Stillman, S. (2003). Instrumental variables and GMM: Estimation and Testing. *The Stata Journal: Promoting Communications on Statistics and Stata*, 3(1):1–31.

Bourricaud, F. (2017). *Poder y sociedad en el Perú contemporáneo*. Instituto de Estudios Peruanos, 3rd edition.

Brülhart, M. (1998). Economic geography, Industry Location and Trade: The Evidence. *The World Economy*, 21(6):775–801.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Chumacero, J. (2012). Industrialización en el perú: 1930–1975. *Pensamiento Crítico*, 17(2):009–026.

Crafts, N. (2004). Market potential in british regions, 1871–1931.

Crafts, N. and Mulatu, A. (2005). What explains the location of industry in Britain, 1871–1931? *Journal of Economic Geography*, 5(4):499–518.

Delgado, A. and Del Pozo, J. M. (2011). Convergencia y ciclos económicos departamentales en el Perú: 1979–2008 - CIES.

Delgado, A. and Rodríguez, G. (2014). Convergencia en las Regiones del Perú: ¿Inclusión o Exclusión en el Crecimiento de la Economía Peruana (1970–2010)? Technical report, Pontificia Universidad Católica del Perú. Departamento de Economía.

Delgado, A. and Rodríguez, G. (2015). Structural Breaks and Convergence in the Regions of Peru: 1970–2010. *Review of Development Economics*, 19(2).

Dirección Nacional de Estadísticas y Censos (1963). *Primer Censo Nacional Económico*.

Ellison, G. and Glaeser, E. L. (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review*, 89(2):311–316.

Ferreira, P. C., Monge-Naranjo, A., and Pereira, L. T. d. M. (2003). Location, concentration, and performance of economic activity in Brazil. Technical report, Policy Research Working Paper Series.

Geary, F. and Stark, T. (2002). Examining Ireland's Post-Famine Economic Growth Performance. *The Economic Journal*, 112(482):919–935.

Gonzales de Olarte, E. (2016). *Una economía incompleta: Perú 1950–2007 : análisis estructural*. Instituto de Estudios Peruanos and Fondo Editorial Pucp.

Gonzales De Olarte, E., Del, E., Segura, P., and Manuel, J. (2012). Lima, una ciudad policéntrica. un análisis a partir de la localización del empleo. *Investigaciones Regionales - Journal of Regional Research*, (23):29–52.

Gonzales de Olarte, E. and Trelles, J. (2004). Divergencia y convergencia regional en el Perú: 1978–1992. Technical report, Pontificia Universidad Católica del Perú. Departamento de Economía.

Gujarati, D. and Porter, D. (2009). *Basic Econometrics*. McGraw-Hill/Irwin, fifth edition.

Harris, C. D. (1954). The Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers*, 44(4):315.

Herrera Catalán, P. (2014). Assessing the impact of agricultural transport costs on industrial agglomeration *.

Instituto Nacional de Estadística e Informática (1977). *II Censo Nacionales Económicos 1974*.

Instituto Nacional de Estadística e Informática (1982). *Censos Nacionales VIII Población III de Vivienda*.

Instituto Nacional de Estadística e Informática (1995). *Migraciones internas en el Perú*.

Instituto Nacional de Estadística e Informática (1996a). *Número de establecimientos censados según departamento, 1994*.

Instituto Nacional de Estadística e Informática (1996b). *Perú: Distancia entre las principales ciudades del país*.

Instituto Nacional de Estadística e Informática (2010). *Perú: IV Censo Económico Nacional 2008. Principales características de los establecimientos de las industrias manufactureras*.

Instituto Nacional de Estadística e Informática (2016). *Perú: Estructura Empresarial 2024*.

Instituto Nacional de Estadística e Informática (2025). *Perú: Estructura empresarial 2024*.

Instituto Nacional de Planificación (1964). *Sexto Censo Nacional de Población*.

Jaramillo, M. and Huamán, R. (2020). Los sectores no primarios y el mercado interno, 1930–1980. In *La economía peruana entre la gran depresión y el reformismo militar 1930–1980*, pages 217–264. Banco Central de Reserva del Perú; Instituto de Estudios Peruanos.

Jiménez, F., Aguilar, G., and Kapsoli, J. (1999). De la industrialización proteccionista a la desindustrialización neoliberal. Technical report, Pontificia Universidad Católica del Perú / Consorcio de Investigación Económica.

Kim, S. (1995). Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860–1987. *The Quarterly Journal of Economics*, 110(4):881–908.

Kim, S. (1999). Regions, resources, and economic geography: Sources of U.S. regional comparative advantage, 1880–1987. *Regional Science and Urban Economics*, 29(1):1–32.

Klein, A. and Crafts, N. (2012). Making sense of the manufacturing belt: determinants of U.S. industrial location, 1880–1920. *Journal of Economic Geography*, 12(4):775–807.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3):483–499.

Krugman, P., Obstfeld, M., and Melitz, M. (2018). *International Economics: Theory and Policy*. Pearson, 10th edition.

Krugman, P. and Venables, A. J. (1995). Globalization and the Inequality of Nations. *The Quarterly Journal of Economics*, 110(4):857–880.

Lobo Collantes, F. (2022). Promoción industrial en los difíciles años ochenta. Crisis del Banco Industrial del Perú, 1980–1990. *Revista Del Instituto Riva-Agüero*, 7(1):275–320.

Lösch, A. (1954). *The Economics of Location*. Yale University Press.

MacKinnon, D. and Cumbers, A. (2019). *An Introduction to Economic Geography: Globalization, Uneven Development and Place*. Routledge.

Maddison, A. (2003). *The World Economy: Historical Statistics*. Development Centre Studies. OECD.

Martinelli, C. and Vega, M. (2023). Reformas económicas y génesis de la inflación: 1963–1980. In *Historia del Banco Central y la política monetaria de Perú: Tomo I*. Banco Central de Reserva del Perú, primera edition.

Martínez-Galarraga, J. (2012). The determinants of industrial location in Spain, 1856–1929. *Explorations in Economic History*, 49(2):255–275.

Matos Mar, J. (2010). *Las migraciones campesinas y el proceso de urbanización en el Perú*. https://centroderecursos.cultura.pe/es/registrobibliografico/las-migraciones-campesinas-y-el-proceso-de-urbanizaci%C3%B3n-en-el-per%C3%BA.

Midelfart-Knarvik, K. H., Overman, H., Redding, S. J., and Venables, A. J. (2000). The location of European industry. https://ec.europa.eu/economy_finance/publications/pages/publication_summary11139_en.htm.

Midelfart-Knarvik, K. H., Overman, H., and Venables, A. (2001). Comparative advantage and economic geography: estimating the determinants of industrial location in the EU. *European Economic Review*. Working paper version; final publication not found in image.

Milanovic, B. (2016). *Global Inequality*. Harvard University Press.

Ministerio de Hacienda y Comercio (1969). *Anuario estadístico del Perú 1966: Resultados de 1958 a 1966*. https://books.google.com.pe/books?id=5aFPAQAAQMAJ.

Ministerio de Transportes y Comunicaciones (2021). *200 años de historia, 52 años conectando vidas*. https://www.gob.pe/institucion/mtc/informes-publicaciones/2042029-200-anos-de-historia-52-anos-conectando-vidas.

Miranda, J. J., López-Rivera, L. A., Quistberg, D. A., Rosales-Mayor, E., Gianella, C., Paca-Palao, A., Luna, D., and Huicho, L. (2014). Epidemiology of road traffic incidents in peru 1973–2008: incidence, mortality, and fatality. *PloS One*, 9(6).

Missiaia, A. (2019). Market versus endowment: explaining early industrial location in Italy (1871–1911). *Cliometrica*, 13(1):127–161.

Nikolic, S. (2018). Determinants of industrial location: Kingdom of Yugoslavia in the interwar period. *European Review of Economic History*, 22(1):101–133.

Oficina Nacional de Estadística y Censos (1975). *Anuario estadístico de 1969: Resultados de 1967 a 1969.*

Organismo Supervisor de la Inversión en Infraestructura de Transporte de Uso Público (2024). *Anuario Estadístico 2023.*

Osinergmin (2016). La Industria de la electricidad en el Perú. Technical report, Osinergmin.

Palomino, J. and Rodríguez, G. (2019). Peru's Regional Growth and Convergence in 1979–2017: An Empirical Spatial Panel Data Analysis. Technical report, Pontificia Universidad Católica del Perú. Departamento de Economía. Documento de Trabajo.

Pardo y Lavalle, M. (1862). *Estudios sobre la provincia de Jauja.* Impr. de la época.

Pepper, J. (2002). Robust inferences from random clustered samples: an application using data from the panel study of income dynamics. *Economics Letters*, 75(3):341–345.

Pierson, P. (2011). Politics in Time: History, Institutions, and Social Analysis. Technical report, Princeton University Press.

Piketty, T. (2014). *Capital in the Twenty-First Century.* The Belknap Press of Harvard University Press.

Ronsse, S. and Rayp, G. (2015). What determined the location of industry in Belgium, 1896–1961? *Journal of Interdisciplinary History*, 46(3):393–419.

Rosés, J. R. (2003). Why Isn't the Whole of Spain Industrialized? New Economic Geography and Early Industrialization, 1797–1910. *Journal of Economic History*, 63(4):995–1022.

Sanguinetti, P. and Volpe Martincus, C. (2009). Tariffs and manufacturing location in Argentina. *Regional Science and Urban Economics*, 39(2):155–167.

Scheidel, W. (2017). *The Great Leveler: Violence and the History of Inequality from the Stone Age to the Twenty-First Century.* Princeton University Press. In The Princeton Economic History of the Western World Ser.

Seminario, B. (2016). El desarrollo de la economía peruana en la era moderna: precios, población, demanda y producción desde 1700. *Apuntes: Revista de Ciencias Sociales*, (78).

Seminario, B., Zegarra, L., and Palomino, L. (2019). *Evolución del PIB departamental y análisis de la desigualdad regional en el Perú: 1795–2017.* Inter-American Development Bank.

Seminario, B., Zegarra, L., and Palomino, L. (2020). *Datos del PIB departamental y análisis de la desigualdad regional en el Perú (1795–2017).* Inter-American Development Bank.

Slater, D. (1975). Underdevelopment and spatial inequality approaches to the problems of regional planning in the third world. *Progress in Planning*, 4(2):97–167.

Smith, C. T. (1968). Problems of regional development in Peru. *Geography*, 53(3):260–281.

Spolaore, E. and Wacziarg, R. (2013). How deep are the roots of Economic Development? *Journal of Economic Literature*, 51(2):325–369.

Stock, J. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In Andrews, D. W. K., editor, *Identification and Inference for Econometric Models.* Cambridge University Press.

Superintendencia de Banca y Seguros (1983). *Memoria y Estadística 1983.*

Superintendencia de Bancos (1965). *Memoria y Estadísticas 1965.*

Sutton, B., Lindow, G., Serra, M. I., Ramirez, G., and Pazmino, M. F. (2006). Regional Convergence in Latin America. Technical report, International Monetary Fund.

Thorp, R. and Bertram, G. (2020). *Perú: 1890-1977. Crecimiento y políticas en una economía abierta*. Fondo Editorial de la Universidad del Pacífico.

Toro González, D. (2004). The Location Choices in the Colombian Manufacturing Industry, 1980–1999. *Economía & Región*, 1(2).

United Nations (1958). International Standard Classification of All Economic Activities - Rev. 1. Technical report.

United Nations (1968). International Standard Classification of All Economic Activities - Rev. 2. Technical report.

United Nations (1969). Growth of the world's urban and rural population, 1920 - 2000. Technical Report Population Studies, Vol. 44.

Vega-Centeno, M. (1988). Desarrollo industrial y exportaciones industriales. *Economía*, 11(21).

Venables, A. J. (1996). Equilibrium Locations of Vertically Linked Industries. *International Economic Review*, 37(2):359.

Weber, A. (1929). *Theory of Location of Industries.* Russell & Russell.

Wolf, N. (2007). Endowments vs. market potential: What explains the relocation of industry after the Polish reunification in 1918? *Explorations in Economic History*, 44(1):22–42.

Wooldridge, J. (2020). *Introductory Econometrics.* Seventh edition.

Wooldridge, J. M. (2018). *The Econometric Analysis of Crosssection and Panel Data.* The MIT Press.

Zucman, G. (2019). Global wealth inequality. *Annual Review of Economics*, 11:109–138.