Estudios de Economía. Vol. 29 - Nº 1, Junio 2002. Págs. 139-163

AGGLOMERATION ECONOMIES: INFLUENCE ON THE DISTRIBUTION OF FOREIGN INVESTMENT IN CHILE*

MIGUEL ANGEL QUIROGA SUAZO

Abstract

The purpose of this work is to empirically study the importance of agglomeration economies in the selection of geographical zones in which foreign investors are willing to develop their investment projects in Chile.
For this purpose a discrete choice model, the conditional logit model (CLM), first proposed by McFadden (1974), was used in this study and the application of the IIA test proposed by Hausman and McFadden (1984) is reported. Furthermore, the use of a less restrictive model is proposed. The main advantage of this latter model is that it allows the variances of the random components to be different across alternatives, that is to say it is a Heteroscedastic Extreme Value Model (HEVM).
The results suggest that agglomeration economies significantly influence the selection of the geographical zone in which the investment is to be located. The results are robust to different specifications of the model. Furthermore, the IIA assumption seems to be appropriate and the result of the HEVM is not significantly different of the CLM.

Resumen

El objetivo de este trabajo es estudiar empíricamente la importancia de las economías de aglomeración en la elección de la Región en la cual el inversionista está dispuesto a desarrollar su proyecto en Chile.
Para este propósito se utiliza un modelo de elección discreta, el modelo logit condicional (MLC) planteado por McFadden (1974), y se reporta la aplicación del test de IAI propuesto por Hausman y McFadden (1984). Además, se propone la utilización de un modelo menos restrictivo que el anterior. La principal ventaja de este último modelo, conocido como Modelo de Valor Extremo Heterocedástico (MVEH), es que permite que las varianzas del componente aleatorio sean diferentes entre alternativas.
Los resultados encontrados sugieren que las economías de aglomeración influyen significativamente en la elección de la Región en la cual la inversión

* The author is thankful for the comments and questions of participants in the "Conference of the Economic Development, Technology and Human Capital" and in the Seminar of the Department of Economics of the University of Concepción. I am especially grateful of the following people: Iván Araya, Ricardo Bebczuk, Eugenio Bobenrieth, Carlos Chávez, Jorge Dresdner and Carlos Oyarzún for their comments.
□ Department of Economics. University of Concepción. e-mail: mquirog@udec.cl

se localizará. Estos resultados son estables a diferentes especificaciones del modelo. Además, el supuesto de IAI parece ser apropiado y los resultados del MVEH no son significativamente distintos a los que se obtienen con el MLC.

Keywords: *Localization, agglomeration economies, Foreign Direct Investment, Discrete Choice Models*.

JEL classification*: F2, R12, R2.*

1. Introduction

The current weak situation of the Chilean economy is a reflection of widely varying levels of economic growth at the regional level. Some regions have a low rate of productive activity expansion and elevated rates of unemployment. This situation, together with the recognition of the importance of investment as a determining factor of regional income levels and employment opportunities, has caused the majority of the regional governments – and also the state who is giving support to the extreme zones and the more depressed areas of the country with special measures – to be concerned with the design of mechanisms that attract investors to their geographic area.

The economic literature sets out that efforts to stimulate productive activity in certain areas can be effective in the long term if there are positive externalities due to industrial agglomeration. Otherwise, such efforts are useless in the long term because they will not permanently locate investment in these areas since once the mechanisms are either discontinued or emulated by other zones, the productive activity will resume its traditional pattern of localization. Nevertheless, there are no studies in Chile that recognize the importance of agglomeration economies in the localization of investment and even fewer works report factors that determine the localization of investment. The only reference on the matter is a series of interviews of 35 industrialists, the objective being to identify mechanisms that stimulate the placement of firms in regions (Foundation and Development, 1997).

Consequently, the objective of this work is fundamentally empirical. It consists of analyzing the influence of agglomeration economies at the moment of choosing the region of Chile where foreign investment project is to be located. Although several studies have had similar objectives, these works have been carried out in very few countries: The United States (Carlton, 1983; Coughlin *et al.*, 1991, Head *et al*. 1995, 1999), Portugal (Guimarães *et al*., 2000), Puerto Rico (Guimarães *et al*., 1996) and some Asian countries (Kinoshita, 1998); some of these studies have even only concentrated on Japanese investment (Head *et al*. 1995, 1999). The reason for this is probably due to the difficulty of relying on detailed information of foreign investment and because very few countries have records at the regional level that allow the control of factors that determine the localization of firms.

This study is possible in Chile because information exists for both authorized and materialized foreign investment and also because annual surveys are carried out in the manufacturing sector.

From the methodological point of view, this investigation uses an adaptation of the CLM originally set out by McFadden (1974) to study how the characteristics of certain geographic zones or localities influence the localization decisions of the investor. The use of this methodology complies with the fact that the restrictions imposed in the model allow a lower level of computer resources when running the model empirically. This is especially important when the number of alternatives exceeds four. In this case, a less restricted model such as the multinomial probit is very difficult to estimate because it involves solving high-dimensions integrals (Limdep 7,0 1998).

The use of this method requires the assumption that the random components of the benefits of the different alternatives are assumed to be independently and identically distributed, which is known as the IIA assumption (McFadden, 1974; Amemiya, 1981; Maddala, 1983; Green, 1999). This assumption is difficult to intuitively justify and the mentioned empirical works do not report statistical tests that allow verification of its fulfillment[1]. The application of the IIA test proposed by Hausman and McFadden (1984) is reported here. Additionally, the main methodological contribution of this work is to propose the use of a less restricted method, the HEVM, set out almost simultaneously by Bhat (1995) and Allenby and Ginter (1995)[2] [Green, 1999; LIMDEP 7.0]. This model is more general than the Conditional Logit, since it allows the stochastic terms to not be homoscedastically distributed, thus allowing for different cross elasticities between all pairs of alternatives. In this work, the original model has been adapted to allow heteroscedasticity and it can be seen that the results reported for this method are not significantly different from those obtained from the use of the CLM.

The document is organized as follows. In Section 2 some background is presented on the variables that influence the localization of investment. Then the methodology that has been used in the empirical estimation is described in Section 3; a description of the series and variables is includes. In the Section 4 the estimation results and their implications are presented. Finally, the most important conclusions appear in Section 5.

2. Localization of Investment: A Background

It may be possible that there is dissociation between the objective of reallocating productive efforts and the economic theory. If this is the case, it is probable that efforts and resources are being squandered that could be more effectively used to solve other social problems. This is more evident when the following predominant approaches are contrasted.

One point of view indicates that the use of instruments to encourage localization of investment in certain geographic areas is useless in the long term.

---

1 One exception is the work of Guimarães, P. *et al.* (1996) who use a nested logit model and estimated the model in two stages in such as to cluster together the alternatives that are not independent.

2 Bhat (1995) used this method to estimate a demand for different forms of transport while Allenby and Ginter (1995) used it in a marketing context.

This is due to their not generating permanent effects in the localization of investment, since once the mechanisms are discontinued locally or emulated by other zones, the investment will resume its traditional pattern of localization. The location selected by the investors would be determined by the classical theory of comparative advantages. In this case, the geographic distribution of investment and production is explained by the spatial differences in the endowment of productive factors, technologies and preferences. That is, those factors that determine variations in the production costs and relative prices between regions (results that corroborate this affirmation have been obtained in the works of Kinoshita 1998, Hanson 1996, Richardson and Smith 1995).

Another group of work maintains an opposite view. They indicate that the incentives for the localization of investment are effective even in the long term. Krugman (1991, 1992) maintains that this is based on the existence of certain positive externalities that arise with the concentration of economic (urbanization economies) and industrial (localization economies) activity. This allows for the appearance of external economies of scale for firms: what the literature calls agglomeration economies. Behind this approach lies the idea that the localization of certain productive activities in a certain zone, whether it be by historical accident or initial advantages of localization (including incentives), generates permanent effects because it increases the probability that new investors also choose these zones at the time of deciding the geographic location of investment projects. In this case the advantages of location, once obtained, tend to be perpetuated (Krugman 1991, 1992; Either 1982; Junius 1997).

Therefore, from the point of view of economic policy design, resolving which approach to adopt in practice is not a trivial problem since these two approaches lead in very different directions in terms of economic policy. If it could be verified that the forces linked to the classical theory are those that mainly predominate in practice, the elimination of commercial barriers would be the principle element that determines investment localization. In this case it is unlikely that, by means of incentive structures, permanent changes in the investment localization patterns can be obtained, since, in spite of maintaining these instruments indefinitely through time, the tendency to equalization in these incentive structures would lead to the localization patterns reverting to their traditional structure. If on the other hand, the empirical evidence showed that the localization of economic activity has characteristics of a non-ergodic system (Arthur 1986, 1990), it could exhibit remarkable and irreversible differences with small changes in the initial conditions. In this case the advantages of localization once obtained would tend to be perpetuated through time (Haufler and Wooton 1999, Wheeler and Mody 1992). In this case, systems of incentives that promote the installation of firms in certain regions would generate permanent effects that would remain in place, even after these incentives are discontinued or are emulated in other zones.

As mentioned, the presence of agglomeration economies constitutes the central hypothesis of this work. The reasons that have been used to explain the existence of agglomeration economies can be grouped in two: urbanization economies and localization economies (Junius 1997, Henderson 1999, Eberts and McMillen 1999). In urbanization economies, costs diminish as the volume of regional production in the firm’s increases, thus allowing firms to reduce their transport and commercialization costs (Mitra 1999). In these cases, a fundamental role is played by infrastructure (Eberts and McMillen 1999). However, the scope of this work is not centered on the presence of urbanization economies but rather on localization economies. This concept was originally raised by Alfred Marshall. He argued that the observed concentration of industrial localization comes from the existence of external economies of scale that arise when firms are located near others of the same industry. He suggested three explanations to justify the existence of these agglomeration economies: the existence of joint markets of specialized labor, the availability of industry-specific inputs and services, and the ease of inter-firm technological diffusion.

As indicated by Marshall, industrial localization creates a joint market of workers with specialized skills. Dumais *et al*. (1997) found this effect to be significant. There are two benefits for the workers. First, if the demand of labor by the firms is not perfectly correlated, the spatial concentration makes it possible for workers that are dismissed to have a greater probability of finding employment in other firms. Therefore an increase in the number of companies reduces the probability that the worker remains unemployed for a long period of time. Second, the workers will have more incentive to invest in specialized skills because the possibilities of benefiting from them are increased (Rotemberg and Saloner 2000). The firms also benefit from this concentration since it increases the probability that they can find a worker with specific skills when such a worker is required. Also the decrease of risk for the worker increases the supply of specialized labor and reduces the risk-premium that has to be incorporated in the wages.

Another of the arguments put forward is the existence of increasing benefits derived from the co-localization of economic units (Dumais *et al*. 1997). A specific example comes from the regional groups of suppliers of specialized services (Either 1982). A greater number of suppliers (each one providing a different service) allows a greater division of labor in intermediate input markets, allowing the final producer to out-source some of the jobs in which the company is not specialized, therefore lowering the average cost of the product. The users and suppliers of intermediate inputs have incentives to be located close to each other when there is a combination of economies of scale and moderate costs of transport. In this case, agglomeration allows for the reduction of transport costs and makes it possible to generate sufficiently high demand levels that guarantee the effort to produce highly specialized components (Krugman 1992). Henkel *et al*. (2000) showed that the decision of a firm to invest in a certain zone, for diverse reasons (including historical accident), can promote the creation of such specialized services reinforcing the attractiveness that the zone has for other investors.

One of the most sited arguments to justify the existence of positive externalities linked to localization is the diffusion of technology (Ellison and Glaeser 1997, Baldwin and Forslid, 1999). A great amount of useful technical information may flow through the communications between industrialists, designers and engineers, in a variety of industries. Physical proximity facilitates these knowledge flows and allows for a less expensive means of communication (Maurel and Sédilot 1999, Baptist 2000).

In this work, external economies are approached from the traditional perspective initially put forward by Alfred Marshall (Henderson 1999, Junius 1997). For this reason we must include in the estimation some proxy variable for agglomeration economies. The following precautions have been taken when choosing the indicator: to verify that the results are neither sensitive to the indicator selected for measuring industrial agglomeration nor to the year in which the variable is measured.

Furthermore, regional differences in the paid prices for productive factors can also determine differences in production costs and therefore variations in the expected benefits for the firms. Wages, the price of the energy and the land prices have been included in the model as factors that could determine differences in inter-regional production costs. If these variables do influence production costs and these costs are significant different between regions, the coefficients related to these variables should be negative.

It is recognized in the industrial-agglomeration literature that a relative abundance of factors in certain regions could also cause a strong concentration of industrial activity, especially of the industries that intensively use these factors. Nevertheless, unlike what happens in the case of agglomeration caused by spatial externalities, this type of localization source does not generate positive externalities that warrant any public intervention. In fact, the greater relative abundance of certain factors will determine that in those zones the prices of these factors are lower. However the advantage of localization will tend to disappear because, as more activity flows towards that zone, it will elevate the relative price of the most abundant factor and so eliminate the initial advantages of localization.

3. The Econometric Model Structure

3.1. The Choice Model

The foreign investor who has decided to invest in Chile must undertake multiple initiatives. Including in these is the selection of the location of investment. This choice is discrete and involves different phases. In the broadest sense –that which is considered in this work– it consists of deciding in which of the 13 regions (the administrative divisions of the country) to materialize the investment.

It is considered that the investor chooses the Region that yields the highest benefits. The magnitude of those benefits depends on the characteristics of the Region and on how they influence the income and costs of the project.

Hence, an investor will choose Region *i* if the benefits that are obtained from that location are greater than those obtained from any of the alternative locations. That is if:

$$B_i > B_j \qquad \forall \; j \neq i \quad j = 1,..., i ,...13 \tag{1}$$

Subscript *j* represents the thirteen regions, including those that were not chosen by the investor. The benefits of the firm, in any region *j*, include two components: one deterministic and the other random or stochastic. Hence:

$$B_j = (\beta , x_j) + e_j \tag{2}$$

The non-observable stochastic variable, ε_j, gathers all the idiosyncratic elements that are not known to the econometrist. The deterministic component, $_j$, reflects the systematic effect that certain regional attributes generate in the firm's benefits. This last function depends on the vector of regional attributes, x_j, and on the vector of parameters to be estimated, β.

The stochastic nature of the investor's decision means that the probability that the investor will choose Region *i* is equal to the joint probability that the benefits of Region *i* are greater than the benefits that the project would have obtained in any other region. Formally:

$$(3) \qquad P_i = \text{Prob}\,[B_i > B_j] \qquad \forall\ j \neq i$$

Hence:

$$(4) \qquad P_i = \text{Prob}\,[\ _i(\beta, x_i) + \varepsilon_i > \ _j(\beta, x_j) + \varepsilon_j] \qquad \forall\ j \neq i$$

Expressed in another way:

$$(5) \qquad P_i = \text{Prob}\,[\varepsilon_i + \ _i(\beta, x_i) + \ _j(\beta, x_j) \ > \ \varepsilon_j] \qquad \forall\ j \neq i$$

Therefore:

$$(6) \qquad P_i = \int_{-\infty}^{\infty} f(\varepsilon_i) \left[\prod_{\substack{j=1 \\ j \neq i}}^{13} \left(\int_{-\infty}^{\varepsilon_i + \Pi_i - \Pi_j} f\langle \varepsilon_j \rangle \, d\varepsilon_j \right) \right] d\varepsilon_i$$

Where, $f(\bullet)$ is the density function of probability.

3.2. Method of Estimation

Conditional Logit Model[3]

Theoretically, the value of parameter b can be obtained assuming that the errors are normally distributed. The resulting multinomial probit model can accommodate very general error structures. Although this method has the advantage of not imposing many restrictions on the behavior of the investor, it has the disadvantage of requiring numerical integration that makes the running of the model difficult when the number of alternatives is greater than four (Limdep version 7,0 1998)[4]. For this reason, practically all previous works have used the CLM proposed by McFadden (1974), to estimate the vector of these parameters β. This method, despite imposing restrictions in the modeling of individual behavior, has the advantage of simplifying the estimation process.

---

3 The model put forward in this section is based upon the work of Carlton (1983), Coughlin *et al*. (1991), Guimarães *et al*. (1996, 2000) y Head *et al*. (1995, 1999). These works all used the CLM to estimate the variables that influence the localization of foreign investment.

4 Amemiya (1981) mentions that when the number of alternatives is $m+1$, m integrals must be evaluated.

McFadden (1974) showed that if the errors (ε_j) are independent and identically distributed by a log-Weibull (extreme value type I), then the probability that the investor chooses region *i* can be estimated from:

$$(7) \qquad P_i = \int_{-\infty}^{\infty} e^{-\varepsilon_i}\, exp\left(-e^{-\varepsilon_i}\right) \prod_{\substack{j=1 \\ j \neq i}}^{13} exp(-e^{-\varepsilon_i - \Pi_i + \Pi_j})\, d\varepsilon_i$$

Generating a probability function that does not require numerical integration:

$$(8) \qquad P_i = \frac{e^{\Pi_i}}{\sum_{j=1}^{13} e^{\Pi_j}}$$

The parameter β is estimated using the Maximum Likelihood Method.

The main limitation associated with the use of this method comes from the use of the IIA assumption. In this work, the application of this assumption implies that the relative probability of choosing a certain Region over another is independent of the presence of the other twelve regions. That is to say, the relative probability that investors choose Region VIII instead of Region VII, for example, remains the same whether they are choosing between only these two Regions or they are choosing between all thirteen. For this reason, the method is adequate when the regions are different and there are no regions that can be considered as being close substitutes by the investor. Another limitation is that the assumption imposes restrictions on the substitution between alternatives causing the cross elasticities between pairs of alternatives to be the same.

Heteroscedastic Extreme Value Model (HEMV)

Bhat (1995) puts forward two methods –in addition to that which was mentioned at the beginning of the previous section (the Multinomial Probit model)– that could be used to relax the IIA assumption in such a way as to not impose too many rigidities to the substitution patterns between alternatives.

The first maintains the assumption of the errors being identically distributed (typically either normally distributed or as a type-I Extreme Value), but it allows the random components to be correlated, relaxing the independence assumption. Guimarães *et al*. (1996) adopted this method to study industrial localization in Puerto Rico. They considered the errors to be type-I-extreme-value distributed; the model that is obtained is the Nested Logit. The advantage of this method is that it is much simpler than the Multinomial Probit. However, the empirical implementation forces the clustering of alternatives that are considered near substitutes to each other. Then, the investment decision is made in two stages: first a set of regions is chosen and then a Region is chosen that is located within the chosen set. Therefore, the disadvantage associated to the use of this method is that in many cases –including the decision that we are studying– the definition of the sets in which the choices were grouped is not so evident.

The second way of relaxing the IIA assumption consists of maintaining the independence assumption, but to allow the random components to not identically distributed. In this case the variances of the random term may be different between alternatives. This probably happens in the case when the variance of a variable, that is not observed by the econometrist and that does not affect the location decision, is different between alternatives. For example, in the case of the choice of the localization region of investment, if the availability of qualified manual labor is a non-observed variable whose values vary considerably in Region VIII, but little in the Metropolitan Region, ceteris paribus, the random component in Region VIII will have a larger variance than the Metropolitan Region.

Although this method (HEMV) requires more computer resources than the nested logit, it has the advantage of not requiring the a priori establishment of an arbitrary cluster of alternatives. Bhat (1995) used a similar model to estimate the demand for different means of transport[5]. Bhat´s work was carried out generalizing the CLM to allow heteroscedasticity. The model is known as the Heteroscedastic Extreme Value Model.

In this work, this model is used to relax the IIA assumption. The results produced from this estimation are comparable to those obtained using the CLM. additionally, this method is used because it allows for different cross-elasticities between pairs of alternatives.

The HEVM has the same benefit function (2) with two components: one deterministic and the other stochastic. However in this case, the errors are distributed as Type-I Extreme Value with precision parameters θ_j – the scale parameter is $\sigma = 1/\theta_j$, Then

$$F(\varepsilon_j) = exp\left[-exp(\vartheta_j \varepsilon_j)\right] \tag{9}$$

The ε_j are independent, but not identically distributed. The probability that region *i* is chosen is found in equation (3).

Hence:

$$P_i = \int_{-\infty}^{\infty} \prod_{\substack{j=0 \\ j \neq i}}^{12} F\left[\theta_j\left(\Pi_i - \Pi_j + \varepsilon_i\right)\right] \theta_i f\left(\theta_i\, \varepsilon_i\right) d\varepsilon_i \tag{10}$$

Where f(•) is the density function. The probabilities and derivatives must be evaluated numerically, since a closed form of the integral does not exist. Bhat (1995) indicated that these can be adequately approximated using the Gaussian-Laguerre quadrature. If the scale parameters of the random component for all alternatives are equal, the expression of probability in the last equation is transformed into the one from the Conditional Logit.

---

5 Allenby y Ginter (1995) simultaneously put forward a similar model in a marketing investigation.

3.3. Description of the series and variables used in the estimation

As previously mentioned in the previous model, the investor decides to develop his project in the Region that allows him to yield the highest profits. Therefore, he chooses the Region where the difference between income and production costs is maximum. The characteristics of the project are independent of the selected region, therefore the only factors that can determine whether the profits to the investment are different, are the characteristics of each Region. The distinctive elements of each regional reality can influence the benefits of the investment by affecting the income or the costs of the project. The explanatory variables in this work consist of these variables that affect the investment returns and are not common to all the regions.

It is assumed in this work that the deterministic component of the benefit function is a linear combination of the regional attributes that affect the yield of the investment[6]. In this way:

$$\pi_j\left(\beta, x_j\right) = \beta x_j \tag{9}$$

These mentioned attributes are those produced from an indirect benefit function, such as the product price or the factors used in the production process. The productive factors included as variables in the estimation are: labor, energy and land. The cost of capital is not included because we assume that a regional-level integrated capital market exists. Along with these, other variables were included that measure agglomeration economies. Table 1 shows the listing of the variables that were used as proxy variables in the regional attributes vector.

Description of the series used in the estimation

The basic information used in the estimation of the model was provided by the Foreign Investment Committee. The database facilitated by this Committee contains records at the individual level of authorized foreign investment through Law-Decree 600 between the years 1974 and 1998 in the manufacturing industrial sector.

For each investment, the following information is known: the investor's nationality, the sector of industry, the product, the Region where the project was authorized and the amount of investment – both authorized and materialized.

It was not possible to rely on the information of the year in which the investment was authorized. This constitutes a limitation because it is difficult to know the existing economic conditions at the time when the investor had to inform of the Region in which the investment project would be placed. In order to weigh the importance of this problem, multiply estimations were carried out in which the explanatory variables were measured considering different periods. The results showed to be not sensitive to the period in which the variables were measured. The coefficients –especially those linked to the interest variable measuring agglomeration economies– did not show significant changes when varying

[6] Works published in this area follow a similar direction. See: Carlton (1983), Coughlin *et al*. (1991), Guimarães *et al.* (1996, 2000) and Head *et al.* (1995, 1999).

the year in which the explanatory variable was measured or when obtaining the independent variable from an average of several years. This is probably because the period in which the investor carries out the investment evaluation is not clear. For example, it is important for the investor to project the future value of the variables that influence the result of a project, but the period that must be considered in this exercise is not so evident. Sometimes the value of the variables in the past are used to project their evolution into the future but it is difficult to argue that the past values of the variables would continue appearing in the future. Also, the periodic information at the regional level is not available at the moment project evaluation. For this reason, it is difficult for the investor to observe the current behavior of the variables at the time of making the investment decision.

The database originally contained 884 records of investment. In this work, investments of multi-regional character and those classified as non-effective were excluded (investments that at some moment were authorized, but were not materialized). The latter were not considered because in many cases the records of these investments were incomplete. In the end, the database contained 640 records, each one corresponding to an authorized investment project. The regional distribution of these projects appears in the appendix.

The investors face the same set of choices and must choose one of the 13 regions in which the country is administratively divided. Therefore, in this model the dependent variable is discrete. For each investment record this variable takes a value of 1 in the region where the investment was authorized and 0 in the other regions that were not chosen by the investor.

Description of the variables used as regresors

Agglomeration economies

The indicator used is the number of establishments that have been previously located in a certain industry plus one (N+1). Head *et al.* (1995, 1999) and Gimarães et al. (1996) also used this variable as a measurement of agglomeration economies. The constant 1 is added because the investor needs to incorporate the possible effect of the localization decision to the number of companies (Head *et al.* 1999). The coefficient linked to this variable is expected to be positive, since an increase in the value of this variable increases the probability that other investors also choose this region. Since we cannot rely on information of the year that the foreign investment was materialized or authorized, the number of establishments prior to the decision of the investor is not known. It is therefore considered that the number of establishments in some of the periods is representative of what the number of companies that remain in that industry in the region would be.

The results of the National Survey of the Manufacturing Industry (ENIA) in different years were used (1991, 1995-1997) to calculate this variable. As was indicated, the model has been run several times, each time varying the year in which the variable is measured, ceteris paribus, with the intention of verifying that the results are not sensitive to the choice of the reference year. In general, the results of the runs proved to be robust in regards to the chosen year.

TABLE 1

DEFINITION AND EXPECTED IMPACT OF THE EXPLICATORY VARIABLES

(all variables are expressed as natural logarithms[7])

VARIABLE	TYPE P	TYPE A	EXPECTED SIGN	DESCRIPTION
AGGLOMERATION ECONOMIES				
AGLOM			+	Number of establishments in the industry to which the investment belongs + 1. The type of industry is classified in accordance with CIIU to two digits[8]. The number of establishments is measured in the year 1991.
LABOR			+	Number of people employed in the industry to which the investment belongs. The variable is calculated based on the average of the years 1995-1997.
PRODUCTION COSTS				
WAGE		- or +		Average salary in the industry to which the investment belongs. The variable is calculated based on the average of the years 1995-1997.
UNSKILL			- or +	Participation of unskilled manpower in the industry to which the investment belongs. The variable is calculated based on the average of the 1995-1997.
ENER		-		Energy prices in the industry to which the investment belongs. The variable is calculated based on the year of the 1991.
LAND			-	Regional population density in the year 1999.
INCOME				
GDP PC			+	Average regional GDP per capita (1974-1995).
NEIGH			+	Sum of the average GDP per capita in neighboring regions (1974-1995).

P = permanent variable

A = alternative variable

[7] The variable UNSKILL is a exception because it is not expressed as a natural logarithm.

[8] This classification is maintained for the variables LABOR, WAGE, UNSKILL and ENER.

With regard to the industrial classification, we used CIIU classification (2nd revision) at a level of disaggregation of two digits. Some have put forward that this level of disaggregation could tend to generate less reliable results with respect to the importance of agglomeration economies. Concerning this point, Moomaw (1998) compared the results from using a CIIU classification to two digits and one to three digits, and showed that a smaller level of disaggregation does not exaggerate the importance of the agglomeration economies.

Carlton (1983), Coughlin *et al*. (1991) and Gimarães *et al.* (2000) used employment levels to measure these spatial externalities. In line with these works, some of the runs of the model include the total number of people employed per industry to capture agglomeration economies. As in the previous case, the ENIA (2nd revision) was used with a level of disaggregation of two digits

Determinants of production costs

Wage levels have been obtained using the ENIA in different periods[9]. They were calculated by dividing the total value paid per remuneration concept by the number employed. In the studies where wages have been used as an explanatory variable, the results with respect to the significance and the sign of this variable have been different. In Coughlin *et al*. (1991), the wages were a significant variable that were a disincentive for foreign investment; meanwhile Carlton (1983) could not affirm that the coefficient of this variable was not different to zero. Additionally, it has been noticed that the coefficient for wages could be positive when it is reflecting differences in the skill level of labor. Guimarães *et al*. (2000) and Head *et al*. (1999) obtained results that would indicate this. We included the unskilled proportion of total employment as a proxy for the differences in skill. This variable is included when the coefficient for wages is positive. If this variable is a good proxy for the different skill levels, we expect that a greater participation of unskilled manpower decreases the probability of a foreign investor choosing that Region.

The energy prices have been calculated from the ENIA of 1991 using a disaggregation that is based on the two-digit CIIU code (revision 2). In Carlton (1983) energy prices were one of the significant factors that determined the localization of foreign investment; higher prices for electrical energy and natural gas in certain regions discourage the localization of investment in these zones.

Just as in Guimarães *et al*. (1996, 2000), land cost was included as a factor that determines the localization of foreign investment. Following these works, we have used the population density as a proxy measurement for the land value. A higher population density should discourage investment localization. While not being significant in Guimarães *et al*. (2000), this variable was significant in Guimarães *et al*. (1996) and had the expected sign.

Determinants of the product price

The national market being limited, the majority of foreign investment that is carried out in the manufacturing industry looks to satisfy external demand. There-

---

9 The results proved to be robust in regard to the period in which the wages were calculated.

fore, it is mainly geared towards exporting. That is why proximity to large industrial centers should not be justified with the intention of obtaining higher income. Nevertheless, we have included the product per capita so as to include the possibility of the firm making a regional policy of prices discrimination. Higher income should lead to a lower price elasticity of demand and higher prices. Therefore, a higher level of income per capita is expected to increase the probability that the investor chooses that Region. Spatial considerations also suggest income levels in neighboring regions may be important so we also include a variable that groups the regional product per capita with those of the neighboring regions. Carlton (1983) and Head *et al*. (1999) included these types of variables: in the former they turned out to be highly significant with the expected sign, but in the latter the coefficients were not significantly different from zero.

4. Estimation and Report of Results

The importance of industrial agglomeration in the localization decision of foreign investment has been considered using the CLM and the HEVM. The results are reported in Tables 3 and 4 respectively.

4.1, Conditional Logit Model

Even though the MVEH is a more general model, results for the CLM are usually reported. This is mainly for two reasons. First, the CLM has been extensively used in the agglomeration-economies literature which makes it possible to rely on comparison patterns that allow us to verify that there are no peculiarities in the database that cause the results to be very different from those of other countries. In addition, the comparison of the results obtained using the HEVM and the CLM allows us to appreciate the advantages of using a model with less restrictive assumptions.

In order to obtain a consistent estimation of the parameter associated with industrial agglomeration, the model has been controlled by other variables that, according to the theoretical model and to empirical works developed in other countries, influence the choice of a locality.

The parameters of the model were estimated with the maximum likelihood method. The base model was estimated. Additional runs of the model were carried out with the purpose of verifying that the results were not sensitive to the following: the inclusion or exclusion of variables; the method of calculating the variables; and the year in which the variables were calculated. As a result of these runs, in all cases, the parameter associated with the agglomeration economies variable had the expected sign and were significantly different from zero. The obtained parameters are very stable and the results are not sensitive to the specifications of the model.

A selection of the main results appears in Table 2. Five different specifications are appraised. The results obtained when running the base model are reported in column (1). Column (2) is included to show the effect generated in the results by an alternative method of measuring agglomeration economies. Column (3) shows the results obtained when two variables were included to reduce

TABLE 2
RESULTS OF THE CONDITIONAL LOGIT MODEL ESTIMATION

	(1)	(2)	(3)	(4)	(5)
AGLOM	1,221* (14,952)	1,412* (14,914)	1,233* (14,267)		1,294 (31,273)
WAGE	-0,055 (-0,192)	-0,012 (-0,045)	0,469 (1,323)	0,289 (0,991)	
UNSKILL			1,691** (2,107)	4,216* (5,623)	
ENER	0,514 (1,097)	0,761 (1,741)	0,591 (1,238)	-0,574 (-1,449)	
LAND	0,063 (1,072)	0,085 (1,442)	0,056 (0,916)	0,993* (32,156)	
GDP PC	0,472** (2,548)	0,489** (2,712)	0,204 (0,883)	2,090* (9,160)	
NEIGH			0,498** (2,443)	-0,315 (-1,367)	
Log Likelihood Function	-683,24	-697,76	-678,17	-809,24	-688,73
Adjusted Pseudo R^2	0,584	0,575	0,587	0,507	0,580
Likelihood Ratio Test					
[Restricted / Non Restricted]			[(1)/(3)]	[(4)/(3)]	[(5)/(3)]
Likelihood Ratio			10,14*	262,14*	21,12*
Degrees of freedom			2	1	6
Test for IIA					
Regions excluded	III, VII & XI	III, VII & XI	III, VII &XI	III, VII &XI	III, VII & XI
Calculated Chi-squared	6,07	3,73	5,43	31,56	2,41
K degrees of freedom	5	5	7	6	1
Pr(C>c)	0,299	0,590	0,607	0,000	0,121

Note: t- statistics are found in parenthesis.
* Statistically significant at 1% level
** Statistically significant at 5% level
Statistically significant at 10% level

the importance of the IIA assumption. Finally, the last two specifications show the effect of the inclusion or exclusion of the variable that measures agglomeration economies in the explanatory power of the model. In each of these cases, the results derived from the application of the IIA test are reported at the end of each column. Additionally, the application of the likelihood ratio test is reported for specifications (3), (4) and (5).

For specification (1) the results reported are derived from a standard specification that is controlled by the differences in the productive factors costs, differences in the income per capita and the number of firms that are belong to a certain industry. This last variable captures the presence of agglomeration economies (AGLOM). The results confirm the findings of other empirical studies in that the agglomeration variable is significantly different from zero and has the expected sign. That is to say, it confirms the hypothesis that the greater the number of companies operating in a regional industry, the higher the probability is that other investors of the same industry choose that Region.

As in other studies, the coefficient associated with the size of the market, measured through regional GDP, showed the expected sign as well as being significantly different from zero. The coefficient for the GDP per capita was also significantly different from zero and gave the expected sign. The remaining variables, which show the influence of the factor prices of in the firm's production costs, were not significantly different from zero.

In order to verify that the previous conclusions are not sensitive to the method of measuring the agglomeration variable, the model was rerun – this time using the number of employees in the industry as a proxy variable for agglomeration economies. Carlton (1983), Coughlin *et al*. (1991) and Guimarães *et al*. (2000) used this same variable. Column (2) shows the results of this exercise. As can be seen, the results do not change significantly: especially those linked to agglomeration economies.

Even though the choice of the proxy variable for agglomeration economies does not significantly influence the value of the parameters, the number of firms is used as an agglomeration indicator in the following stages. This choice was made because, when measuring the agglomeration variable in this way, the explanatory power of the regression equation is greater.

When using the CLM, it has been assumed that the stochastic component of the benefits is independently distributed. This being the case, this method imposes a restriction in the modeling of individual behavior to simplify the estimation process. Nevertheless, not complying with this assumption can cause some consistency problems in the estimation. A possible reason for this is that the GDP per capita in the adjacent regions influences the localization decisions. For example, it is possible that many firms are located in the V or VI Region due to their proximity to the Metropolitan Region. In this case, the exclusion of this variable could cause a bias in the estimation of the agglomeration economies coefficient.

The mentioned problem may occur if the skilled workforce is not homogeneously distributed between regions. In this case, it is possible that the agglomeration of firms in a certain zone is the product of a greater availability of skilled or unskilled labor in that zone. The exclusion of this variable could bias the coefficient that measures agglomeration economies since the stochastic term will be correlated with agglomeration economies. In order to reduce these problems, the model was rerun, this time including the sum of the GDP per capita of the neighboring regions and the participation of the unskilled workforce of the industry as explanatory variables. The results of this exercise appear in column (3). The parameters associated to both variables have the expected sign and are significantly different from zero. The application of the likelihood ratio test – reported in the column (3) – on comparison with that which appears in column

(1), shows that both specifications are significantly different when considering a level of significance of 1%. This implies that the average income of the population in the adjacent regions and the participation of skilled labor are not variables that influence the investor's localization decision.

The Independence of Irrelevant Alternatives assumption

Despite including the latest two variables to limit the possibility that the IIA assumption is not fulfilled, it is convenient to apply some specification test that provides an idea of how serious the problems continue to be when the model is run based on this assumption. The fulfillment of this assumption implies that eliminating some of the irrelevant alternatives will not alter the relative probabilities between the remaining choices. That is to say, the relative probability between a pair of alternatives is specified without considering the nature of a third option (Amemiya 1981). In 1984, Hausman and McFadden published a test to consider the validity of this assumption. It is to be noted that none of the referenced works report results derived from such a test.

Hausman and McFadden (1984) indicate that if a set of alternatives is truly irrelevant, completely omitting it does not produce any systematic changes in the estimators of the parameters. In this case, the elimination of the alternatives would generate inefficiency but not inconsistency. If however, the hypothesis is not accepted, meaning that the remaining relative probabilities are not independent of these alternatives, the estimators obtained by eliminating these alternatives are inconsistent. The alternative statistic proposed by Hausman and McFadden (1984) is:

$$\chi^2 = \left(\hat{\beta}_r - \hat{\beta}_{nr}\right)' \left[\hat{V}_r - \hat{V}_{nr}\right]^{-1} \left(\hat{\beta}_r - \hat{\beta}_{nr}\right) \tag{10}$$

Subscript *r*, in β and *V*, indicates that the parameter vector and the asymptotic covariance matrices, respectively, were considered with the restricted set of choices. The subscript *nr*, in β and *V*, indicates that the parameter vector and the asymptotic covariance matrices were considered including all the possible alternatives. The asymptotic distribution of the statistic is chi-square distributed with K degrees of freedom (Green, 1999). The IIA test was applied to all the specifications of the model that are reported in Table 2. The results of the test when Regions III, VII and XI are excluded appear in the last rows of Table 2. For these Regions, we can affirm that the IIA assumption is fulfilled for all specifications reported in the table. For the others, the exclusion of one or more regions generates a matrix that is not positive definitive, thus making the application of the test impossible. The results obtained by the conditional logit in these cases should therefore be observed with care.

Importance of Agglomeration Economies

The last two columns of Table 2 present two specifications that show the importance of the industrial agglomeration variable when explaining the localization of the foreign investment in the manufacturing industry. In order to verify the importance of agglomeration economies in the explanatory power of the model, the empirical model was rerun without considering agglomeration economies. The result of this exercise appears in column (4). As can be seen, the explanatory power of the model diminishes considerably; the likelihood ratio test reported in the column shows that both specifications are significantly different. What is more, a rather simpler model – such as the one reported in the last column containing only industrial agglomeration as an independent variable – was found to have a greater explanatory power: which is reflected in a higher value of the log-likelihood function.

It should be noted that the estimations of the other parameters are not very stable. The exclusion of the agglomeration variable caused the magnitude, the sign and the significance of several parameters to change.

Therefore, it is appraised that the central variable of this work, the one that measures agglomeration economies, is always statistically significant (with a level of significance less than 1%) and its coefficient has the expected sign. Hence, we can conclude that agglomeration economies do affect the probability of choosing a certain Region. In this case, an increase in the number of firms of a certain industry in a Region will increase the probability that new foreign investors with projects related to that industry will choose that Region at the time of carrying out their investment projects.

4.2. Heteroscedastic Extreme Value Model

The MVEH is used to decrease the number of restrictions for the estimation of the parameters. The output of this estimation is reported in Table 3. It can be seen in this table that the parameter associated with agglomeration economies, being significantly different from zero, is related directly to the localization of investment. The other parameter that turns out to be significantly different from zero is the participation of the unskilled workforce, which shows a positive sign.

A specification similar to that reported in column (3) of Table 2 was used in the estimation of the MVEH. This allowed the results to be compared using a less restricted model such as the CLM or a flexible one such as the MVEH. The results from these models are not significantly different. The application of the likelihood ratio test, comparing this result with the one obtained using the CLM, shows that both specifications are not significantly different, considering a 10% level of significance.

The previous conclusion is reinforced when analyzing the estimated value of the scale parameters that are reported in the same table. Although all of them are significantly different from zero, it cannot be affirmed –with the exception of Regions V and VI– that they are significantly different from 1: this being the value assigned for the scale parameter of the Metropolitan Region.

A comparison analysis of the signs and the significance of the parameters using both models shows that the difference only comes from the influence of the GDP per capita of adjacent regions, which is no longer significantly different from zero when the HEVM is used. One of the advantages of the HEVM is that it allows for cross elasticities between different alternatives. In this case, although the great majority of the scale parameters are not statistically different, the differences in the parameter levels could determine a greater variation of the elasticities between alternatives.

TABLE 3
ESTIMATE RESULTS FOR THE HETEROSCEDASTIC EXTREME VALUE MODEL (HEVM)

	Variable	Coefficient Value	t- statistic
	AGLOM	0,422*	5,405
	WAGE	0,219	1,346
	UNSKILL	0,985*	2,974
	ENER	0,065	0,318
	LAND	0,094	1,317
	GDP PC	0,643	1,590
	NEIGH	-0,248	-0,608
			Standard Deviation
	θ R1	2,139	0,978
	θ R2	3,915	2,065
	θ R3	1,727	0,557
	θ R4	2,221	0,774
	θ R5	3,177*	1,094
	θ R6	2,372*	0,592
	θ R7	1,785	0,506
	θ R8	4,167	2,062
	θ R9	1,651	0,523
	θ R10	1,561	0,466
	θ R11	1,331	0,443
	θ R12	2,879	1,039
θ R13 (fixed Parameter)		1,000	
Log Likelihood Function		-687,55	
Pseudo-adjusted R^2		0,581	
Likelihood Ratio Test [Restricted / Non Restricted]			
Likelihood Ratio		º18,76	
Degrees of Freedom		12	
Chi-squared **		21,03	

* Statistically significant at 1% level
Statistically significant at 5% level
Statistically significant at 10% level
θ Rj = Scale Parameter “estimated” for Region j

4.3. Elasticities

The estimated parameters can be used to analyze the effect that a variation in the parameter *m* in Región *j* [x_j (m)] has on the probability that the investor choices location *i* [P_i]. In this model that change corresponds to:

$$\frac{\partial P_i}{\partial x_j(m)} = \left(1 - P_j\right) P_i \,\beta(m) \tag{11}$$

The independent variables are expressed as logarithms in this study. Therefore, it is convenient to modify expression (11) to reflect the elasticities. This results in:

$$(12) \qquad \phi_{ij}(m) = \frac{\partial P_i}{\partial x_j(m)} \, \frac{x_j(m)}{P_i} = \frac{\partial P_i}{\partial \log x_j(m)} \, \frac{1}{P_i} = \left(1 - P_j\right) \beta(m)$$

This elasticity can be interpreted as the percentage variation of the probability that a foreign investment is localized in a certain Region *i*, due to a 1% change in value of one of the independent variables in Region *j*.

This investigation has focused on the importance of industrial agglomeration in the localization of foreign investment. Hence, Table 4 reports the elasticities values when the number of firms located in a Region *j* is changed, ceteris paribus. The specification for column (3) of Table 2 was used to calculate these elasticities. The headings of the columns in Table 4 give the Region *j* that experiences a 1% increase in the number of firms; the rows give the resulting change in the probability of investment localization in the Region *i* due to this increase in Region *j*. For example, an increase of 1% in the number of firms located in Region VIII would cause an increase of 3.46% in the probability of future foreign-investment projects being located in that same Region, thereby diminishing by 0.36% the probability that these projects are located in any of the other regions.

Table 5 shows the exact same calculations except using the HEVM reported in Table 3. It can be seen that the MVEH allows a more flexible pattern of elasticities.

TABLE 4

THE PROBABILITY ELASTICITY OF THE NUMBER OF FIRMS IN AN INDUSTRY
CONDITIONAL LOGIT MODEL

		Change in the number of firms in a Region												
		I	II	III	IV	V	M	VI	VII	VIII	IX	X	XI	XII
Effect on the probability that an investment is placed in these Regions	**I**	**2,262**	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	II	-0,085	**2,242**	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	III	-0,085	-0,069	**0,781**	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	IV	-0,085	-0,069	-0,010	**1,427**	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	V	-0,085	-0,069	-0,010	-0,032	**3,157**	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	M	-0,085	-0,069	-0,010	-0,032	-0,205	**1,643**	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	VI	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	**1,718**	-0,068	-0,301	-0,032	-0,131	-0,005	-0,030
	VII	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	**2,146**	-0,301	-0,032	-0,131	-0,005	-0,030
	VIII	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	**3,425**	-0,032	-0,131	-0,005	-0,030
	IX	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	**1,484**	-0,131	-0,005	-0,030
	X	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	**2,090**	-0,005	-0,030
	XI	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	**0,548**	-0,030
	XII	-0,085	-0,069	-0,010	-0,032	-0,205	-4,384	-0,051	-0,068	-0,301	-0,032	-0,131	-0,005	**1,581**

TABLE 5
PROBABILITY ELASTICITY OF THE NUMBER OF FIRMS IN AN INDUSTRY
HETEROSCEDASTIC EXTREME VALUE MODEL

		Change in the number of firms in a Region												
		I	**II**	**III**	**IV**	**V**	**M**	**VI**	**VII**	**VIII**	**IX**	**X**	**XI**	**XII**
Effect on the probability that an investment is placed in these Regions	**I**	**0,075**	0,066	0,023	0,042	0,095	0,162	0,51	0,064	0,107	0,044	0,064	0,016	0,046
	II	-0,114	**2,849**	-0,028	-0,023	-0,615	-3,324	-0,124	-0,069	-0,582	-0,032	-0,105	-0,006	-0,100
	III	-0,089	-0,209	**0,781**	-0,023	-0,368	-2,466	-0,095	-0,064	-0,380	-0,028	-0,115	-0,007	-0,091
	IV	-0,124	-0,142	-0,014	**1,717**	-0,509	-2,403	-0,113	-0,109	-0,588	-0,036	-0,202	-0,012	-0,141
	V	-0,001	0,000	0,000	0,000	**0,007**	-2,967	0,000	0,000	0,000	0,023	-0,101	-0,007	-0,077
	M	0,010	-0,033	-0,004	-0,005	-0,083	**0,834**	-0,018	-0,023	-0,093	-0,009	-0,032	-0,002	-0,016
	VI	-0,076	-0,145	-0,013	-0,022	-0,224	-2,180	**1,530**	-0,076	-0,478	-0,029	-0,124	-0,007	-0,095
	VII	-0,053	-0,063	-0,007	-0,015	-0,171	-1,951	-0,051	**1,551**	-0,444	-0,030	-0,098	-0,006	-0,053
	VIII	-0,092	-0,159	-0,014	-0,028	-0,356	-2,588	-0,103	-0,152	**3,840**	-0,064	-0,181	-0,011	-0,129
	IX	0,453	0,476	0,169	0,359	0,763	-2,064	-0,065	-0,085	-0,506	**1,226**	-0,120	-0,007	-0,065
	X	-0,047	-0,041	-0,007	-0,014	-0,129	-1,581	-0,045	-0,056	-0,287	-0,023	**1,203**	-0,006	-0,051
	XI	-0,063	-0,054	-0,011	-0,020	-0,180	-1,758	-0,056	-0,084	-0,350	-0,034	-0,134	**0,391**	-0,061
	XII	-0,102	-0,146	-0,014	-0,032	-0,325	-2,322	-0,114	-0,102	-0,704	-0,038	-0,171	-0,010	**1,687**

5. Conclusions

The study of the factors that determine the regional distribution of foreign investment in the manufacturing sector allows us to appreciate the importance of agglomeration economies. Within the scope of this study, we have observed that the estimator of the parameter associated with the agglomeration variable is always statistically significant (with a level of significance less than 1%) and the coefficient has expected sign. Furthermore, the incorporation of the agglomeration variable significantly increase the explanatory power of the model.

There is therefore evidence that agglomeration economies have established the regional distribution of foreign investment in the manufacturing industry of Chile. That is to say, an increase in the number of industry-specific firms in a certain Region will cause an increase in the probability that new foreign investors, with projects of that same industry, choose that Region at the time of carried out an investment project.

This result has interesting applications in economic policy: especially those policies that promote the decentralization of economic activity at the national level. Although it does not suggest which instruments should be used, this result confirms that government policies that manage to geographically redirect productive investment would be effective because they would cause an increase in the probability that new investors also choose those regions at the moment of locating their investments. Consequently, the location of economic activity would constitute a non-ergodic system and more active government policies oriented to stimulating the localization of productive activity in certain regions would be effective.

One must recognize that, as with all empirical investigations, this study has some limitations. As is mostly the case, an important limitation is the availability of information. It would have been desirable to know the exact year in which investment projects were authorized or finalized in order to better control the factors that then determined the choice of location. Furthermore, variables may exist that influence the localization decision, which have not been taken into consideration. In this case, if these excluded variables are correlated with the number of firms of an industry, we will obtain a biased measurement of the agglomeration effect. This could happen due to not including variables that capture differences in the endowment of productive factors or in transportation costs.

Due to these obvious limitations, many other regressions were run that have not been reported in this work. The objective was to verify that these results were not sensitive to the assumptions and simplifications that had to be made. In general, we can affirm that the results are established neither by the method used in the estimation nor by the exclusion or inclusion of variables. These results are even insensitive to the method of measuring agglomeration economies, and so, when agglomeration economies are measured on the basis of employment, the results do not change.

Another aspect dealt with in this work is the IIA assumption on which the CLM is based. Although the use of this assumption facilitates the estimation of the model, it is not very easy to justify from an intuitive point of view. The works in which the CLM is used to estimate the importance of different factors that influence investment localization do not report the application of any specification test to verify the fulfillment of this assumption. We followed two paths in this study: the test of Hausman and McFadden was used (1984) to verify that the IIA assumption was fulfilled and, we estimated a less restricted model that relaxed the IIA assumption by allowing heteroscedasticity. The use of the test of Hausman and McFadden (1984) allowed us to verify that the IIA assumption was fulfilled. Moreover, the estimation of the more general model, HEVM, and its later comparison with the CLM, showed that both models are not significantly different, The HEVM, however, has the advantage of permitting cross-elasticities between different alternatives.

6. References

Allenby, G. and J. Ginter (1995). The effects of in-store displays and feature advertising on consideration sets. *International Journal of Research in Marketing*, 12, 67-80.

Amemiya, T. (1981). Qualitative Response Models: A survey. *Journal of Economic Literature* 19, 1483-1536.

Arthur, B. (1986). Industry location patterns and the importance of history. Stanford University, Center for Economic Policy Research, paper nº 84.

Arthur, B. (1990). Positive feedbacks in the economy. *Scientific American* 262, 92-99.

Baldwin, R. and R. Forslid (1999). The core-periphery model and endogenous growth: stabilising and de-stabilising integration. *National Bureau of Economic Research*, NBER Working Paper 6899.

Baptista, R. (2000). Do innovations diffuse faster within geographical clusters? *International Journal of Industrial Organization* 18, 515-535.

Bhat, C. (1995). A Heteroscedastic Extreme Value Model of intercity mode choice. Working Paper Department of Civil Engineering, University of Massachusetts, Amherst.

Carlton, D. (1983). The location and employment choice of new firms: an econometric model with discrete and continuous endogenous variables. *The Review of Economics and Statistics* 65, 440-449.

Coughlin, C., J. Terza, and V. Arromdee (1991). State characteristics and the location of foreign direct investment within the United States. *The Review of Economics and Statistics* 73, 675-683.

Dumais, G., G. Ellison, and E. Glaeser (1997). Geographic concentration as a dynamic process. *National Bureau of Economic Research*, NBER Working Paper 6270.

Eberts, R. W. and D. P. McMillen (1999). Agglomeration Economies and Urban Public Infrastructure. En: *Handbook of Regional and Urban Economics* (Ed.: Cheshire and Mills), Vol. 3, Applied Urban Economics, North Holland, pp. 1455-1495.

Either, W. (1982). National and international returns to scale in the modern theory of international trade. *American Economics Review* 72, 389-405.

Ellison, G. And E. Glaeser (1997). Geographic Concentration in U. S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105, 5: 889-927.

Fundación Región y Desarrollo (1997). Mecanismos de estímulo para la localización de empresas en regiones. http://www.fundacionryd.cl/proyec/mieco.htm

Green, W. (1999). Análisis Econométrico. 3ª edición. Prentice Hall.

Guimarães, P., R. Rolfe and D. Woodward (1996). Regional Incentives and Industrial Location in Puerto Rico. Working Paper, Center for International Business Education and Research, U.S. Department of Education.

Guimarães, P., O. Figueiredo and D. Woodward (2000). Agglomeration and the Location of Foreign Direct Investment in Portugal. *Journal of Urban Economics* 47, 115-135.

Hanson, G. (1996). Localization Economies, Vertical Organization, and Trade. *The American Economic Review*, 86, 5, 1266-1278.

Haufler, A. and I. Wooton (1999). Country Size and Tax Competition for Foreign Direct Investment. *Journal of Public Economics*, 71, 121-139.

Hausman, J. and D. McFadden (1984). A Specification Test for the Multinomial Logit Model. *Econometrica* 52, 1219-1240.

Head, C. K., J. C. Ries, D. L. Swenson (1995). Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investment in the United States. *Journal of International Economics* 38, 223-247.

Head, C. K., J. C. Ries, D. L. Swenson (1999). Attracting Foreign Manufacturing: Investment Promotion and Agglomeration. *Regional Science and Urban Economics* 29, 197-218.

Henderson, V. (1999). Marshall's Scale Economies. *National Bureau of Economic Research*, NBER Working Paper 7358.

Henkel, J., K. Stahl, and U. Walz (2000). Coalition building in a spatial economy. *Journal of Urban Economics* 47, 136-163.

Hines, J. R. Jr. (1996). Altered states: taxes and the location of foreign direct investment in America. *American Economic Review*, 86, 5, 1076-1094.

Junius, K. (1997). Economies of scale: A survey of the empirical literature. Kiel Institute of World Economics, Kiel Working Paper 813.

Kinoshita, Yuko (1998). Firm Size and Determinants of Foreign Direct Investment. Working Paper CERGE-EI. Prague, Czech Republic.

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

Krugman, P. (1992). Geografía y Comercio. Antoni Bosch editor.

LIMDEP Version 7.0 (1998). User‘s Manual. Econometric Software, Inc.

Maddala, G. S. (1983). Limited–depend and qualitative variables in econometrics. Cambridge University Press.

Maurel, F. and B. Sédillot (1999). A measure of the geographic concentration in French manufacturing industries. *Regional Science and Urban Economics* 29, 575-604.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. En *Frontiers in Econometrics* (Ed: P. Zarembka). Academic Press, New York, pp. 105-142.

Mitra, A. (1999). Agglomeration economies as manifested in technical efficiency at the firm level. *Journal of Urban Economics* 45, 490-500.

Moomaw, R. (1998). Agglomeration economies: Are they exaggerated by industrial aggregation? *Regional Science and Urban Economics* 28, 199-211.

Richardson, J. D. and P. Smith (1995). Sectoral Growth Across U.S. States: Factor Content, Linkages, and Trade. *National Bureau of Economic Research,* NBER Working Paper 5094.

Rotemberg, J. and G. Saloner (2000). Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade. *Regional Science and Urban Economics* 30, 373-404.

Wheeler, D. and A. Mody (1992). International investment location decisions: The case of U. S. Firms. *Journal of International Economics* 33, 57-76.

APPENDIX

FOREIGN INVESTMENT PROJECTS
AUTHORIZED IN THE INDUSTRIAL MANUFACTURING SECTOR YEARS 1974-1998
AND DESCRIPTION OF THE INDEPENDENT VARIABLES OF ESTIMATION

Región	Nº of Projects	AGLOM		WAGE		UNSKILL		ENER		LAND		GDP PC		NEIGH	
		Ave.	S.D.	Ave.	S.D.	Ave.	S.D.	Ave.	S.D.	Ave.	S.D.	Ave.	S.D.	Ave.	S.D.
I	13	2,577	0,598	1,176	0,209	0,687	0,088	3,365	0,146	1,894	0,000	5,926	0,000	6,487	0,000
II	11	2,216	0,292	1,701	0,348	0,503	0,071	3,245	0,325	1,300	0,000	6,487	0,000	5,818	0,000
III	5	1,507	0,228	1,182	0,587	0,682	0,168	3,340	0,090	1,275	0,000	5,696	0,000	6,046	0,000
IV	3	2,292	1,384	0,815	0,206	0,784	0,016	3,279	0,202	2,642	0,000	5,240	0,000	5,626	0,000
V	21	2,893	0,985	1,394	0,356	0,663	0,075	3,231	0,150	4,545	0,000	5,551	0,000	5,675	0,000
VI	15	2,747	0,767	1,293	0,429	0,659	0,184	3,324	0,099	3,861	0,000	5,680	0,000	5,504	0,000
VII	14	2,755	1,096	1,095	0,397	0,654	0,203	3,145	0,206	3,399	0,000	5,167	0,000	5,552	0,000
VIII	28	3,280	1,064	1,356	0,529	0,707	0,161	3,294	0,061	3,945	0,000	5,405	0,000	4,963	0,000
IX	6	1,561	1,478	0,921	0,400	0,785	0,135	3,302	0,085	3,302	0,000	4,706	0,000	5,225	0,000
X	29	3,088	1,239	0,904	0,028	0,745	0,144	3,325	0,056	2,752	0,000	5,006	0,000	5,141	0,000
XI	7	1,222	0,747	1,211	0,496	0,670	0,300	3,335	0,130	-0,148	0,000	5,443	0,000	6,101	0,000
XII	12	2,433	0,841	1,600	0,473	0,648	0,193	3,307	0,088	0,168	0,000	6,611	0,000	5,443	0,000
R.M.	476	4,966	0,695	1,379	0,220	0,625	0,047	3,274	0,130	5,967	0,000	5,716	0,000	5,593	0,000