# Determinants of the location choices in sugar industry of Iran: using the logit & probit model

*Determinanty alokačního výběru v cukrovarském průmyslu v Íránu: využití logit a probit modelu*

Mehdi MOGADDAM[1], Soleyman IRANZADEH[2], Hossein BEVRANI[3]

[a]*State University of Baku, Baku, Azerbaijan and Islamic Azad University, Bonab, Iran*
[b]*Islamic Azad University, Tabriz, Iran*
[c]*University of Tabriz, Tarbiz, Iran*

**Abstract:** The main objective of this research is to offer a model for determining the most proper location for investment in sugar industry. This model classifies the possible options for investment, taking into account several variables. The orientation of this research is an applied orientation and its objective is an explorative one. Therefore, no hypothesis is proposed to prove it. The required data for determining the parameters of the model is extracted from statistical yearbooks of the Iran Ministry of Agriculture during 1996 to 2006. According to the results obtained from this research, if new investments in sugar industry are carried out based on the offered classifications and pattern, the profitability will definitely increase.

**Key words:** location model, logit model, probit model, sugar industry

**Abstrakt:** Hlavním cílem prezentovaného výzkumu je nabídnout model pro determinaci nejvhodnější lokality pro investice v cukrovarském průmyslu. Model klasifikuje možné volby investování, přičemž zohledňuje několik proměnných. Výzkum je orientován na praktickou aplikaci a jeho cíle jsou explorativní, není tudíž navržena žádná hypotéza, jež by byla prokázána nebo vyvrácena. Potřebné údaje pro stanovení parametrů modelu byla získána ze statistických ročenek Ministerstva zemědělství Iránu za roky 1996–2006. Výsledky výzkumu ukazují, že pokud jsou nové investice do cukrovarského průmyslu alokovány podle navrhované klasifikace a vzorce, jejich rentabilita se výrazně zvýší.

**Klíčová slova:** alokační model, model logit, model probit, cukrovarský průmysl

So far, different models have been offered for determining a proper location for this industry. Each of these models is designed for specific environmental conditions and different variables are considered the most proper location in accordance with those conditions.

Decisions for determining the locations in the Iran sugar industry are made traditionally and some variables are considered separately. So, the location selected for investment might not be the best possible one.

This research has tried to solve this problem and to design and offer a comprehensive model for determining locations in this industry, which can consider many variables as integral variables for selecting the optimal location.

The location theory has a rich history of scientific research dating back to the later part of the $19^{th}$ and early part of the 20th centuries. Its beginnings are found in the seminal work of Weber (1929). A renewed interest took hold in the middle 1900s in the work of Hoover (1948), Isard (1948), McLaughlin and Robock (1949) and Greenhut (1956). Common themes in this literature were aimed at empirically testing the theoretical determinants for choosing the industrial location. Moreover, the location determinants of manufacturing investment are likely to evolve as the composition of the manufacturing industry changes. Blair states that the determinants of location choices change as the conditions of production change. In fact, the investigation of the location determinants changing over time, the purpose of this section is to review different models.

## THE RELATED EMPIRICAL LITERATURE

The location of economic activity represents a logical and testable case of the firm behaviour (Guimaraes et al. 2004). According to Porter, the decisions relating to the selection and acquisition of property for the location of a business entity can have a substantial impact on the firm's ability to establish and maintain a competitive advantage (Mazzarol and Choo 2003).

The determinants of industrial location should be well established. We should know a lot about the relative importance of basic economic factors such as the cultivated area, the production rate of raw materials, the unemployed persons ready to work, the distance between factory and selling place of products, the distance between factory and purchasing place of raw materials and the total population of the city.

Yet the results of the vast location empirical literature vary widely. Moreover, the basic questions keep getting recast in different models. What is the real efficacy of the above mentioned variables on location? Almost invariably, the motivation for more empirical research is that these and other major questions remain unanswered (for example, Foster 1977; Leitham et al. 2000; Forslid et al. 2001; Chakravorty 2003; Disdier and Mayer 2004). Evidently, a systematic approach to the industrial location modelling has not been found.

The conditional Logit model is often used in the literature on industrial location. The first applications of this type of model for studying industrial location were from Carlton in1979 and Bartik in1985 (Carod 2005).

More recent studies are those by Disdier and Mayer (2004); Guimaraes et al. (2004); Lambert et al. (2006); Devereux et al. (2007); Arauzo-Carod et al. (2009).

Guimaraes et al. (2004) designed the random utility maximization-based conditional logit model (CLM) which serves as the principal method for the applied research on the industrial location decisions. Studies that implement this methodology, however, confront several problems, notably the disadvantages of the underlying Independence of Irrelevant Alternatives (IIA) assumption. This study shows that by taking an advantage of the equivalent relation between the CLM and Poisson regression likelihood functions, one can more effectively control the potential IIA violation in the complex choice scenarios where the decision maker confronts a large number of narrowly defined spatial alternatives.

Disdier and Mayer (2004) investigate the determinants of the location choices of French multinational firms in Eastern and Western Europe. They find important differences between the two regions of Europe regarding these determinants. Agglomeration effects are less strong in Central and Eastern European countries (CEECs) than in the European Union (EU) countries. Location decisions are influenced significantly and positively by the institutional quality of the host country. Disdier and Mayer also investigate whether investors consider Western Europe and Eastern Europe as two distinct groups of potential host countries. They confirm the relevance of an East–West structure in the country location decision and show that this relevance decreases as the transition process advances in the CEE countries.

Lambert et al. (2006) tested for and incorporated spatial processes using the geographically weighted regression (GWR) and the Poisson regressions when looking at location factors of new manufacturing firms. A distance decay function was used which applies the geo- statistical concepts, the direct approach for modelling spatial dependence. Lambert et al. (2006), in looking at food manufacturing investment, use a spatial probit model to estimate a spatial lag model to account for the spillover effects. Spatial dependence is routinely modelled as an additional covariate in the form of a spatially lagged dependent variable $Wy$, or in the error structure where $E[\varepsilon_i \varepsilon_j] \neq 0$. The first is referred to as the spatial lag model and is utilized when the importance is granted to the presence of spatial interaction. Spatial dependence in the error term can take many forms via the spatial error model and is commonly referred to as the nuisance dependence (Anselin 2003). This specification is appropriate when trying to correct the spatial autocorrelation. A spatial error model can be expressed as:

$$Y = X\beta + \mu$$

$$\mu = \lambda W\mu + \varepsilon \quad (1)$$

where $y$ is a vector $(1 \times N)$ of observations on the dependent variable, $X$ an $K N \times$ matrix of observations on the explanatory variables given in the equation (2), and μ an error term that follows a spatial autoregressive (SAR) specification with the autoregressive coefficient λ. In the spatial auto regression, the vector of errors is expressed as a sum of the vector of random terms ($\varepsilon$) and a so-called spatially lagged error, $W\mu$. Formally, a spatial lag model is given by

$$y = \rho Wy + X\beta + \varepsilon \quad (2)$$

where ρ is a spatial autoregressive coefficient often referred to as the spatial correlation coefficient, $\varepsilon$ is a vector of error terms; $Wy$ is a spatial lag operator which is a weighted average of $y$ variables at the neighbouring locations (Anselin 2003). Neighbouring criteria determine the structure of $W$ which is routinely based off of contiguity (queen or rook) or distance criteria. The weights in $W$ are usually row-standardized, so that elements sum to one. Equations 1 and 2 are most commonly estimated with the maximum likelihood. However, the spatial lag model expressed in the reduced form shows that $Wy$ is correlated with the error term,

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \quad (3)$$

Under this specification, an investment decision in a given state is connected to all other investment decisions by the spatial multiplier $(I - \rho W)^{-1}$ and the error term (Anselin 2003). Equations 1 and 2 are often estimated via maximum likelihood. The spatial econometric literature has identified the advantages of the maximum likelihood estimation (MLE) in the presence of the spatial dependence (e.g. Anselin et al. 1996; Elhorst 2003). Kelejian and Prucha (1999) provide an alternative estimation procedure of spatial process models via the generalized method of moments (GMM). They point to at least two potential advantages of this approach: (1) relaxing the assumption of normality of error terms, and (2) less computationally burdensome compared to the MLE.

Devereux et al. (2007) examine whether the discretionary government grants influence where domestic and multinational firms locate new plants, and how the presence of agglomeration externalities interacts with these policy instruments. They find that a region's existing industrial structure has an effect on the location of new entrants. Grants have a small effect in attracting plants to the specific geographic areas, but importantly, they find that firms are less responsive to government subsidies in areas where there are fewer existing plants in their industry. This suggests that these subsidies are less effective in influencing the firms' location decisions in the face of the countervailing co-location benefits.

Arauzo-Carod et al. (2009) find that the basic analytical framework has remained essentially unaltered since the early contributions of the early 1980s, while, in contrast, there have been advances in the quality of the data (more firm and plant level information, geographical disaggregation, panel structure, etc.) and, to a lesser extent, the econometric modelling. They also identify certain determinants (neoclassical and institutional factors) that tend to provide largely consistent results across the reviewed studies.

## THE RESEARCH MODEL

If there are ***n*** measurable variables effective on location in the real world which cause difference in the desirability of locations, we can write the function of "Location Desirability" that the employer is faced with as follows:

$$Ui = Ui\,(Mi) = Zi\,(Mi) + \varepsilon_i\,(Mi) \quad i = 1, 2, ..., k \quad (4)$$

$i$ is the number of locations being investigated in this equation (number of cities) and $Zi$ is defined as follows:

$$Zi = Zi\,(Mi_1, Mi_2, ..., Mi_n) \quad (5)$$

$M_{in}$ shows the $n^{th}$ measurable characteristic in location of $i^{th}$, so the industry will be settled in a location with the desirability (profitability) higher than the other locations. In other words, the $i^{th}$ option is selected when in comparison with the other option like $j$, it has more desirability or profitability. That is:

$$\begin{aligned} &Zi \neq Zj \;:\; Ui\,(Zi\,)\rangle Uj\,(Zj\,) \\ &\forall j \neq i \end{aligned} \quad (6)$$

Or the probability of selecting the $i^{th}$ option for settlement of the industry is:

$$Pi = P\,(\varepsilon_j - \varepsilon_i < Zi - Zj) \quad (7)$$

Logit & Probit is the most famous random model for selecting a special location and if we use $\chi^2$ (Chi-square) distribution in it and it is assumed that the random part ($\varepsilon_i$) has logistic distribution; the probability of distribution for $Zi$ will be calculated as follows:

$$F(Zi) = \frac{e^{Zi}}{1 + e^{Zi}} \quad (8)$$

In other words, the probability of selecting the $i^{th}$ location among K locations is completed as follows and the options are ranked after calculating the probabilities.

$$Pi = \frac{\exp(Zi)}{\sum_{i=1}^{k} \exp(Zi)} \quad (9)$$

## THE PROPOSED MODEL FOR LOCATING SUGAR INDUSTRY IN IRAN

We will reach the following estimable final model by expanding a discontinuous selection model and defining the effective variables on locating in sugar industry on this account. Locating sugar industry in 30 cities of Iran was studied by the consideration of six variables effective in such undertaking as:

$$Y = a_1 X_1 + a_2 X_2 + ... + a_6 X_6 + u \quad (10)$$

Table 1. Results of the model (dependent variable $Y$, $n$ = 240)

| Independent variables | Logit(1) | Probit(1) | Logit(2) | Probit(2) | Logit(3) | Probit(3) | Logit(4) | Probit(4) | Logit(5) | Probit(5) |
|---|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | 0/001139 (0/353) | 0/00701 (0/299) | 0/00118 (0/115) | 0/000707 (0/094) | 0/00131 (0/077) | 0/00078 (0/057) | 0/00139 (0/035) | 0/000811 (0/029) | /001488 (0/021) | /000853 (/015) |
| $X_2$ | 1/28 (0/962) | 1/75 (0/99) | – | – | – | – | – | – | – | – |
| $X_3$ | 0/000277 (0/738) | 0/000142 (0/744) | 0/000285 (0/729) | 0/000143 (0/73) | 3/31 (0/87) | 1/35 (0/91) | – | – | – | – |
| $X_4$ | 0/0141 (0/293) | 0/00862 (0/241) | 0/014 (0/282) | 0/0086 (0/23) | 0/0153 (0/187) | 0/00912 (0/16) | 0/0164 (0/14) | 0/0095 (0/11) | 0/031136 (0/0088) | 0/0170 (0/0029) |
| $X_5$ | 0/0035 (0/481) | 0/84 (0/529) | 0/0349 (0/48) | 0/184 (0/52) | 0/0344 (0/45) | 0/0188 (0/49) | 0/033 (0/22) | 0/0190 (0/23) | – | – |
| $X_6$ | 0/00555 (0/772) | 0/00285 (0/772) | 0/00577 (0/762) | 0/00577 (0/76) | – | – | – | – | – | – |

In this model, the dependant variable is a dummy variable that chooses (1) for the locations with sugar factories and (0) for locations without a sugar factory.

The independent variables of this method include: $X_1$ (cultivated area), $X_2$ (production rate of raw materials), $X_3$ (unemployed persons ready to work), $X_4$ (distance between the factory and selling place of products), $X_5$ (distance between the factory and purchasing place of raw materials) and $X_6$ (total population of the city).

It is expected that all the variables with the positive sign will appear in this model except $X_4$ and $X_5$.

## DATA ANALYSIS

The data related to independent variables of the suggested model are extracted from statistical yearbooks of the Iran Ministry of Agriculture for the years 1996–2006.

After entering the data in the suggested Logit & Probit model and solving it in Eviews software (version 3.1), the summary of the obtained results are shown in Table 1.

The figures in this table show the coefficient of the variables. The figures in the parentheses show the probability of error of the variable and if they are less than sig < 0.05, it means the significance of the variable; otherwise, it will mean the insignificance of the variable.

According to these results, as the coefficients became significant in the two last columns of the table and the probability of error is less than 0.05 ($P < 0.05$), we can calculate the probability of selecting thirty cities using the following equation:

$$Pi = \frac{\exp(\widehat{Y})}{\sum_{i=1}^{30} \exp(\widehat{Yi})} \tag{11}$$

According to the obtained results, where the coefficient of $X_1$ has become significant, other variables will remain fixed and if this variable (cultivated area of sugar beet) increases up to 1% in a location, the chance of choosing that location for the establishment of a sugar factory will increase up to 0.001488%.

Thus, as other variables remain fixed, if $X_4$ increases up to 1%, the chance of choosing that location for

Table 2. Probability of 10 cities to be selected with high priorities in 1996, 2001, 2006

| Year / prob. | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1996 | Miandoab | Orumiyeh | Naghadeh | Khoy | Mahabad | Shahindej | Tabriz | Salmas | Azarshahr | Bostanabad |
| prob. | 96.84 | 2.284 | 24.024 | 0.1979 | 0.159 | 0.0145 | 0..934 | 0..524 | 0..220 | 0.170 |
| 2001 | Naghadeh | Miandoab | Orumiyeh | Khoy | Mahabad | Salmas | Oskou | Oshnavieh | Azarshahr | Bostanabad |
| prob. | 77.22 | 12.85 | 6.6 | 1.12 | 0..958 | 0.177 | 0.125 | 0.108 | 0.821 | 0.0.57 |
| 2006 | Naghadeh | Miandoab | Orumiyeh | Mahabad | Tabriz | Piranshahr | Chaldran | Bokan | Oskou | Oshnavieh |
| prob. | 77.2 | 20.58 | 1.13 | 0..399 | 0.146 | 0.11 | 0.095 | 0.057 | 0.039 | 0.028 |

prob. = probability

Table 3. Ranking of locations based on the probability of selection in 1996, 2001, 2006

| Rank | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1996 | Miandoab | Orumiyeh | Naghadeh | Khoy | Mahabad | Shahindej | Tabriz | Salmas | Azarshahr | Bostanabad |
| 2001 | Naghadeh | Miandoab | Orumiyeh | Khoy | Mahabad | Salmas | Oskou | Oshnavieh | Azarshahr | Bostanabad |
| 2006 | Naghadeh | Miandoab | Orumiyeh | Mahabad | Tabriz | Piranshahr | Chaldran | Bokan | Oskou | Oshnavieh |

Table 4. The optimal pattern of locating sugar industry

| Rank | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| City | Naghadeh | Miandoab | Orumieh | Khoy | Mahabad |

establishment of a sugar factory will increase up to 0.03113%.

Considering the coefficients of $X_1$ and $X_4$ variables and 1% effect of each, we can conclude that $X_4$ (distance between the factory and selling place of products) in comparison with other independent variables effective on locating is more important and considerably so than economic activities.

As the main objective of this research is to rank the locations in terms of the probability of selection, so ranking of the locations up to 10 priorities is shown in Table 2 according to the results obtained from solving the suggested model and calculating the probabilities (prob.) from the offered equation.

Table 3 indicates the ranking of the above 10 cities, the selection of the probabilities of which is shown in Table 2.

However, to achieve the desired results through referring to the Table 3 and comparing the ranks of cities, the first five cities most likely to profit are Naghadeh, Miandoab, Orumieh, Khoy and Mahabad respectively, which had a higher rank than the other cities. Accordingly, the optimal pattern of sugar locating industry will be as indicated in the Table 4 and Figure 1.

## CONCLUSION

According to the results obtained from this research, if new investments in sugar industry, like an establishment of factory, an increasing of the capacity and modernization are carried out according to the submitted ranking and pattern, the desirability (profitability) will definitely increase. The first option for investment in this pattern is Naghadeh and in case of any unforeseen obstacles like social and political obstacles for the execution of the project, Miandoab, Orumiyeh and etc. are considered for the execution of this project.


Figure 1. The locations of the selected cities with ranks on the real map

This research suggests a relatively new and complete method for industrial location that involves several independent variables. We hope that the productivity of industries will increase by completing and adding other variables and applying this model in different industries.

## REFERENCES


Anselin L. (2003): Spatial externalities, spatial multipliers, and spatial econometrics. International Regional Science Review, *26*: 153–66; DOI: 10.1177/0160017602250972

Anselin L., Bera A., Florax R.J.G.M., Yoon M. (1996): Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics, *26*: 77–104; DOI: 10.1076/0166-0462(95)02111-6

Arauzo-Carod J.-M., Liviano-Solis D., Manjón-Antolín M. (2009): Empirical studies in industrial location: an assessment of their methods and results. Journal of Regional Science, *49*: 1–32; DOI: 10.1111/j.1467-9787.2009.00625.x

Carod J. (2005): Determinants of Industrial Location: An Application for Catalan Municipalities. Paper of Regional Science, 84: 105–120, Blackwell Publishing.

Chakravorty S. (2003): Capital source and the location of industrial Investment: A tale of divergence from post-reform India. Journal of International Development, *15*: 365–383; DOI: 10.1002/jid.976

Devereux M., Griffith R., Simpson H. (2007): Firm location decisions, regional grants and agglomeration externalities. Journal of Public Economics, 91: 413–435; DOI: 10.1016/j.jpubeco.2006.12.002

Disdier A., Mayer T. (2004): How different is Eastern Europe? Structure and determinants of location choices by french firms in Eastern and Western Europe. Journal of Comparative Economics, *32*: 280–296; DOI: 10.1016/j.jce.2004.02.004

Elhorst J.P. (2003): Specification and estimation of spatial panel data models. International Regional Science Review, *26*: 244–268; DOI: 10.1177/0160017603253791

Forslida R., Haalandb J., Knarvikb K. (2002): A U-shaped Europe? A simulation study of industrial location. Journal of International Economics, *57*: 273–297; PII: S0022-1996(01)00155-6

Foster R. (1977): Economic and quality of life factors in industrial location decisions. Social Indicators Research, *4*: 247–265.

Greenhut M.L. (1956): Plant Location in Theory and in Practice. University of North Carolina Press, Chapel Hill; ISBN: 13:9780313233388.

Guimaraes P., Figueiredo O., Woodward D. (2004): Industrial location modeling: Extending the random utility framework. Journal of Regional Science, *44*: 1–20; DOI: 10.1111/j.1085-9489.2004.00325.X

Hoover E.M. (1948): The Location of Economic Activity. McGraw-Hill, New York.

Isard W. (1948): Some locational factors in the iron and steel industry since the early nineteenth century. Journal of Political Economy, *56*: 203–217.

Kelejian H., Prucha I. (2004): Estimation of simultaneous systems of spatially interrelated cross sectional equations. Journal of Econometrics, *118*: 27–50; DOI: 10.1016/S0304-4076(03)00133-7

Lambert D.M., Garret M.I., McNamara K.T. (2006): An application of spatial poisson models to manufacturing investment location analysis. Journal of Agricultural and Applied Economics, *38*: 105–121. Available at http://www.econpapers.repec.org/article/agsjoaaec/43752.htm

Leitham S., McQuaid R., Nelson J. (2000): The influence of transport on industrial location choice: a stated preference experiment. Transportation Research, Part A, *34*: 515–535; PII: S0965-8564(99)00030-0

Mazzarol T., Choo S. (2007): A study of the factors influencing the operating location decisions of small firms. Property Management, *21*: 190–208; DOI: 10.1108/02637470310478918

McLaughlin G.E., Robock S. (1949): Why Industry Moves South. NPA Committee of the South. Report No. 3, Kingsport Press.

Weber A. (1929): Theory of the Location of Industries (translated by Carl J. Friedrich from Weber's 1909 book). University of Chicago Press, Chicago.


Arrived on 21st May 2009

*Contact address:*

Hossein Bevrani, University of Tabriz, Faculty of Mathematical Sciences, Department of Statistics, Tabriz, Iran
e-mail: bevrani@tabrizu.ac.ir