131-20

ITESM

CAMPUS CIUDAD DE MEXICO

INSTITUTO TECNOLOGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

ITESM Campus Ciudad de México-Universidad de Texas en Austin
Programa Doctoral

ACTA DE EXAMEN DOCTORAL

Candidato: *Dídimo Antonio Dewar Valdelamar*

Título de la Tesis: *"Measuring DEA efficiency for the Human Development Index Achieved by Mexican States"*

Fecha de examen: *Abril 24, 1997* **Hora:** *16:00 hrs.*

Asesor:

Nombre: *Dr. Genaro Gutiérrez*

Firma:

Primer Lector:

Nombre: *Dr. David Muñoz Negrón*

Firma:

Segundo Lector:

Nombre: *Dr. David Romero Vargas*

Firma:

Examinador y Director del Programa Doctoral:

Nombre: *Dr. Carlos M. Urzúa Macías*

Firma: C. M. Urz

Decisión: *Aprobado* *Rechazado*

Dr. Javier Morales González
Director
División de Graduados e Investigación

ITESM

CAMPUS CIUDAD DE MEXICO

INSTITUTO TECNOLOGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY

We certify that in the city of Austin, Texas, on April 24, of 1997, the student:

Dídimo Antonio Dewar Valdelamar

has defended his dissertation:

"Measuring DEA efficiency for the Human Development Index Achieved by Mexican States"

that has been presented as the final requirement for the degree of:

Doctor en Administración

Based on the evidence presented by the thesis research and this oral examination, the committee chaired by Dr. Carlos M. Urzúa Macías has taken the following resolution.

Approved

Dr. Genaro J. Gutiérrez
Principal Advisor

Dr. David Muñoz Negrón
Reader

Dr. David Romero Vargas
Reader

C. M. Urzúa

Dr. Carlos M. Urzúa Macías
Examinator and
Director of Doctoral Program

Dr. Javier Morales González
Dean of Graduate Studies
Mexico City Campus

I dedicate this work to my parents Celia and Horacio as a signal of gratitude
for their love, care and support

To my wife Sandra, for being such a wonderful human being

To my beloved sons Diego and David

To my brother Alex and my sisters
Damaris and Nelda

Measuring DEA Efficiency For The Human Development Index Achieved By Mexican States

by

Didimo Antonio Dewar Valdelamar (B.S., MBA)

Doctoral Dissertation

Presented to

Instituto Tecnologico y de Estudios Superiores de Monterrey

in Partial Fulfillment of the Requirements
for the degree of

Doctor of Philosophy

April 1997

2
7
3

Acknowledgements

I hereby express my deepest acknowledgement to all the people who helped me in concluding this dissertation. In first place God for giving me the opportunity to continue although critical moments in this very long process. Special acknowledgement to Dr. Genaro Gutierrez, the chairman of this dissertation, a guide and friend that supported and encourage me to continue, thank you very much for all your observations. To Dr. David Muñoz, and Dr. David Romero, members of my dissertation committee for their comments specially in the final part of this dissertation process. I am deeply indebted with all the rest of professors that were part of this joint effort between ITESM, Campus Ciudad de Mexico and the University of Texas at Austin.

I recognized the great effort of ITESM to assemble this program and the great support offered all the way, from the begining to the end by Dr. Cesar Morales President of ITESM South Zone, Ing. Arturo Soltero who was the ITESM General Director when the program began, Ing. Jorge Ivan Diaz, Dr. Enrique Zepeda and Dr. Alfredo Jacobo General Directors of my Campus, ITESM Campus Morelos, as well as Ing. Francisco Becerril who was Director of the Graduate and Research Division and Lic. Dora Emilia Gonzalez Director of Academic Development of Campus Morelos.

I also express my deepest gratitude to my wife Sandra Viscarra and to my two sons David and Diego Dewar to who I took away many weeks of family time, but who also enjoyed the process and made good friends specially in summer time when we went to Austin. I sweared something, finishing the dissertation, and I fulfilled it, it was complicated but not impossible. Thank you very much for your support and confidence in me mother, Celia Valdelamar de Dewar and father Horacio Dewar, I took a long time but I did not let you down.

I would like to thank to each and everyone of my Ph. D. program generation for their advise and support and friendship specially to Mario De Marchis, Agustin Buendia with who we traveled each Thursday to Mexico City and back to Cuernavaca on Saturdays in each single week of the first part of the program and to Macario Schettino for all his advise, support and humor.

Finally I would also like to thank Supera (Programa de Superacion Academica) who is sponsored by ANUIES (Asociacion Nacional de Universidades e Institutos de Educacion Superior) for its economic support.

ABSTRACT

Measuring DEA Efficiency For The Human Development Index Achieved

By Mexican States

by

Dídimo Antonio Dewar Valdelamar, B.S Industrial and Systems Engineering, MBA

Supervising Professor: Genaro J. Gutiérrez

This dissertation studies the use of DEA (Data Envelopment Analysis) as tool for evaluating the technical efficiency with which Mexican states obtain their respective HDI (Human Development Index). The HDI was developed by the UN (United Nations) to evaluate the level of development of the countries. Two DEA models with same inputs and different outputs were run and evaluated to distinguish their capacity of discriminating between efficient and inefficient states. In one model the output was the HDI and in the other model the outputs were the components of the HDI. To run the DEA models the HDI was calculated first for each of the states and an analysis of the results was discussed. In second place an Optimal Stratification procedure was done for the Mexican states using the the socioeconomical variables that are components of the HDI in order to compare the compatibility of the classification of the states with this technique and the HDI technique. Finally the DEA models were run for determining states efficiency, reference sets for the inefficient states, calculating the optimal level of inputs and outputs for the inefficient sets and sensitivity analysis to suggest adequate social expenditure for Mexican states.

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1 AN OVERVIEW OF THE COUNTRY 1

1.2 EFFICIENCY 9

1.3 ITS IMPORTANCE 13

1.4 RELATION WITH PRODUCTIVITY 13

1.5 MOTIVATION 16

1.6 STATEMENT OF THE RESEARCH 16

1.7 ORGANIZATION OF THE DISSERTATION 17

2. THE UN METHODOLOGY 19

2.1 INTRODUCTION 19

2.2 HUMAN DEVELOPMENT INDEX MODEL 20

3. THE DEA TECHNIQUE 25

3.1 BACKGROUND 25

3.2 THE MODEL 26

3.3 THE DIFFERENT TYPES OF EFFICIENCY 30

3.4 CHARACTERISTICS 32

3.5 DEA MODELS 33

3.5.1 Additive Model 33

3.5.2 Multiplicative Model 34

3.5.3 Banker, Charnes and Cooper (BCC) Model 35

3.5.3.1 Banker, Charnes and Cooper (BCC) Input Orientation Model 36

3.5.3.2 Banker, Charnes and Cooper (BCC) Output Orientation Model 37

3.5.4 Charnes, Cooper and Rhodes (CCR) Model 38

3.6 SENSITIVITY ANALYSIS 39

3.7 LITERATURE REVIEW 41

3.7.1 History 41

3.7.2 Review of Research Applying DEA 41

3.8 LIMITATIONS 50

4. HUMAN DEVELOPMENT INDEX (HDI) OF MEXICAN STATES 51

4.1 Variables 52
4.2 Sources 54
4.3 Results 55
4.4 Discussion 62

5. **OPTIMAL STRATIFICATION OF MEXICAN STATES 68**

5.1 Introduction 68
5.2 Stratification 68
5.3 Optimum Statification in Multivariate Sampling 69
5.4 Results 71
5.5 Discussion 73

6. **DATA ENVELOPMENT ANALYSIS 75**

6.1 Variables 76
6.2 Sources 78
6.3 Results 80
6.4 Discussion 94

7. **CONCLUSIONS 99**

7.1 Summary Contributions and Conclusions 99
7.2 Limitations 102
7.3 Directions for Future Research 103

8. **BIBLIOGRAPHY 104**

9. **VITA 108**

TABLES LIST

TABLE 3-1INPUTS AND OUTPUTS VARIABLES OF CHARNES, COOPER AND LI, 1989 42
TABLE 3-2 INPUTS AND OUTPUT VARIBLES OF BANNISTER AND STOLP, 1995 43
TABLE 3-3 INPUTS AND OUTPUTS VARIABLES OF SINUANY-STERN ET. AL 1994 45
TABLE 3-4 INPUT AND OUTPUT VARIABLES OF CHARNES, COOPER AND RHODES, 1981 46
TABLE 3-5 INPUT AND OUTPUT VARIABLES OF GOLANY, PHILLIPS AND ROUSSEAU, 1993 49
TABLE 4-1 RATIO OF EDUCATIONAL ACHIEVEMENT FOR POPULATION OF 16 YEARS AND OLDER. 56
TABLE 4-2 POPULATION OF 15 YEARS AND OLDER LITERACY RATE 58
TABLE 4-3 LIFE EXPECTANCY, GDP FOR 1988 AND GDP PER CAPITA 59
TABLE 4-4 HUMAN DEVELOPMENT INDEX PER STATE 61
TABLE 5-1STATES STRATIFICATION BASED ON SOCIOECONOMICAL LEVEL. 72
TABLE 5-2 STATES STRATIFICATION BASED ON EDUCATIONAL LEVEL. 73
TABLE 6-1 INPUT VARIABLES (PER CAPITA TERMS) 77
TABLE 6-2 OUTPUTS VARIABLES (PER CAPITA MEASUREMENTS AND HDI). 79
TABLE 6-3 DMU'S ϕ AND EFFICIENCY VALUES. 81
TABLE 6-4 REFERENCE SETS OF BCC PER CAPITA MODEL. 82
TABLE 6-5 CONTINUATION OF REFERENCE SETS FOR BCC PER CAPITA MODEL. 83
TABLE 6-6 REFERENCE SETS OF BCC HDI MODEL. 83
TABLE 6-7 VALUES OF INPUTS AND OUTPUTS OF BCC PER CAPITA MODEL IF DMU'S WERE EFFICIENT. 84
TABLE 6-8 VALUES OF INPUTS AND OUTPUT OF THE BCC HDI MODEL IF DMU'S WERE EFFICIENT 85
TABLE 6-9 INPUTS DUAL VARIABLES AND SUSTITUTION RATES FOR THE BCC PER CAPITA MODEL 86
TABLE 6-10 INPUTS AND OUTPUT DUAL VARIABLES AND SUBSTITUTION RATE FOR THE BCC HDI MODEL.88
TABLE 6-11 OUTPUT DUAL VARIABLES AND SUBSTITUTION RATES FOR THE BCC PER CAPITA MODEL.... 89
TABLE 6-12 SUBSTITUTION RATES BETWEEN OUTPUTS AND INPUTS BCC PER CAPITA MODEL. 91
TABLE 6-13 CONTINUATION OF SUBSTITUTION RATES BETWEEN OUTPUTS AND INPUTS FOR THE BCC PER CAPITA MODEL 92
TABLE 6-14 SUBSTITUTION RATES BETWEEN OUTPUT AND INPUTS OF THE BCC HDI MODEL 93

FIGURES LIST

FIGURE 3-1 GRAPHIC DESCRIBING MEASUREMENT OF EFFICIENCIES. 31

FIGURE 3-2 ENVELOPMENT SURFACE FOR THE ADDITIVE MODEL. 34

FIGURE 3-3 COBB-DOUGLAS ENVELOPMENT SURFACE 35

FIGURE 3-4 ENVELOPMENT SURFACE FOR THE INPUT-ORIENTED BCC MODEL. 37

FIGURE 3-5 ENVELOPMENT SURFACE FOR THE OUTPUT-ORIENTED BCC MODEL. 38

FIGURE 3-6 ENVELOPMENT SURFACE FOR THE INPUT-ORIENTED CCR MODEL. 39

1. Introduction

1.1 An Overview of the Country

Mexico, the northernmost latin american country in the continent, is composed of 31 states and a Federal District, which is the capital and seat of the federal government. Mexico is a federal republic in which, according to the Mexican constitution, each of the 31 states is free and sovereign. It is a country with many natural resources, including oil and natural gas. It also has a huge variety of climates which permits the cultivation of many types of crops, although this is obviously subject to the availability of water in the regions.

The 31 states or federal entities plus the Federal District names are: Aguascalientes, Baja California, Baja California Sur, Campeche, Coahuila, Colima, Chiapas, Chihuahua, Federal District or better known as Mexico City, Durango, State of Mexico, Guanajuato, Guerrero, Hidalgo, Jalisco, Michoacan, Morelos, Nayarit, Nuevo León, Oaxaca, Puebla, Queretaro, Quintana Roo, San Luis Potosi, Sinaloa, Sonora, Tabasco, Tamaulipas, Tlaxcala, Veracruz, Yucatan and Zacatecas.

Aguascalientes is a state in Central Mexico; the state elevation ranges between 900 to 3050 meters above sea level. The capital city of this state has the same name as the state Aguascalientes; the numerous thermal springs that are in the region give name to the state. It is a fertile agricultural region, much of which is irrigated. Crops produced in the region include fruit, corn and grapes, which are used by the state's brandy industry. Livestock raising is an important economic activity. Mining activities include zinc, silver and gold.

Baja California is a state that is part of a peninsula with the same name. Its capital city is Mexicali. The state has a very hot and arid climate; the average annual rainfall in the peninsula is 254 milimeters. Most of its population is clustered in three cities near the U.S. border: Tijuana, Mexicali and Ensenada. Many maquiladoras can be found in this state; the maquiladora industry is the engine for the economic activity in this state. An important winery industry can also be found here.

Baja California Sur became a state in the year of 1974. Formerly it was a territory. Its capital city is La Paz. This state is also part of the peninsula with the same name. As its name states, it is the southern part of the peninsula. The most important economic activity is tourism. The tourism activity has grown rapidly because of the completion of roads.

Campeche is a southeastern state of Mexico and it is located on the Yucatan Peninsula. The city of Campeche is the state capital. It became a state in 1867. In the south of the state, a humid tropical rain forest can be found; an important economic activity is forestry. In the north, which is arid, agriculture and stock raising is an important activity. Campeche is a port city, and it lies in the bay of Campeche; commercial fishing is also a very important economic activity.

Coahuila is a northern state, and Saltillo is its capital city. Most of the state is an arid plateau. Coahuila and Texas were united from 1824 to 1836; it became a Mexican state in 1868. Coahuila is traditionally a stock raising area. Irrigation has made agriculture possible, and grains, cotton, and grapes are grown. The state is Mexico's leading coal producer and copper, iron, lead and silver are also mined in the state.

Colima is a western central state located on the Pacific Coast. Colima is its capital city. The principal economic activities are: raising livestock and cultivating lemons, rice, coffee, cotton and maize. Copper, iron lead and gold are mined.

Chiapas is the southernmost state in Mexico. It is situated on the Pacific Coast; its capital city is Tuxtla Gutierrez. The state is mostly mountainous and the maximum elevation in the state is 3140 meters above sea level. The economy is based upon stock raising, rubber, cacao and coffee cultivation. Natural gas and petroleum resources are exploited. Its population is primarily composed of Mayan Indians. Chiapas has a very beautiful natural areas which help attract tourism, an economic activity that generates important income for the state.

Chihuahua is the largest state of Mexico, it is located in the northern part of the country. It covers 244,938 square kilometers. The city of Chihuahua is its capital city. The principal economic activity is mining, where iron, lead, zinc, silver and gold are exploited. Cattle raising, cotton, apples and nuts cultivation are also very important.

The Federal District is the capital of the country. It is considered the leading industrial and cultural center of Mexico and also one of the world's fastest-growing urban areas. The factories that are located in Mexico City use almost half of the electricity generated and account for more than half of its industrial output, meanwhile employing around 60% of the manufacturing workforce. The manufacturing industries include: food processing, textiles, pharmaceuticals and automobile assembly. The Federal District is the hub of Mexico's transportation network. Tourism is also a significant contributor to the economy.

Durango is a state located in north central Mexico; its capital city is also named Durango. The major economic activity is livestock raising on the semi-arid plateau.

In the valley irrigated by the Rio Nazas, different products are grown, such as corn, cotton, wheat, tobacco and vegetables. Mining in the mountains is another important economic activity where gold, silver, coal, iron among other minerals are mined. In the towns, there are textile, timber, metallurgy and food processing activities. In 1823, when Mexico became independent, Durango became a state.

Mexico State, is a state that is located in the center of the country, and it envelops the Federal District. The capital of the state is Toluca. Its most important economic activity is agriculture, especially dairy farming, but mining and tourism are also important economic activities. Many manufacturing industries have been established in this state, where chemicals, pharmaceuticals and automobile assembly are the most noteworthy.

Guanajuato is an important state in central Mexico, and as in several other states, its capital city has the same name as the state. Mexico´s independence war began in this state and it became a state in 1824. Mining is a very important activity and gold, silver, lead and copper are mined. Agricultural activities are also important and the products grown are: corn, barley, beans and wheat. Industries are concentrated in urban areas such as the cities of Leon and Celaya. Industrial activities relevant for the state economy are: textile manufacturing, food processing, petrochemicals production and tanning. The city of Leon is known as the shoe capital of the country. Tourism is also an important activity for the state´s economy.

Guerrero, along Mexico´s Pacific coast, is a southwestern state. Its state capital is Chilpancingo de los Bravos. Guerrero became a state in 1849. Its most important economic activity is tourism, in its important cities as Acapulco, Ixtapa and Taxco de Alarcon. Other economic activities in the state are agriculture, forestry and mining;

the most important products of the state are: corn, cotton, tobacco, coffee, gold, lead, iron, silver and mercury.

Hidalgo is a central state of Mexico; Pachuca de Soto is its capital. The basis of its economy wealth and industrial activities are related to mining, especially mining of silver.

Jalisco is a western central state of Mexico; its capital city is Guadalajara. This state has a very diversified economy where tourism, industries and agriculture are important activities. The cities of Guadalajara and Puerto Vallarta are the most attractive for tourists. The principal agricultural products are: corn, wheat, beans, tobacco and citrus fruits. Stock raising is also an important economic activity. Among its industries, tanning and metal mechanics are important.

Michoacan is a western central state of Mexico along the Pacific Ocean coast, its capital city is Morelia. Agriculture, mining and livestock raising account for more than 50% of the state's economy. Its elevation range is from sea level up to 3050 meters.

Morelos is the second smallest state of Mexico; it lies south of Mexico City. Its capital city is Cuernavaca. Morelos became a state in 1869 and was named for the country's national hero, Jose Maria Morelos y Pavon. Morelos is an agricultural region; the most important products are sugarcane, rice, onions, fruit and vegetables. Another very important economic activity is tourism. Some industrial activity can be found in its state capital, the most important being auto assembly done by the Japanese automaker Nissan.

Nayarit is a very mountainous state in western cental Mexico; it borders the Pacific Ocean. Its capital city is Tepic. Agriculture is its main economic activity, where corn, beans, coffee, tobacco, sugarcane, and cotton are important. Raising livestock and mining activities are also important for the economy of the state.

Nuevo Leon is a border state in northeastern Mexico. Its capital city is Monterrey. This is a state with a well balanced economy. It became a state in 1824. In the agricultural sector, cotton, wheat and citrus fruits are the most predominant. Livestock raising is also important in the state. The state is the home of several of the most important industrial groups. Among its most important industrial products are: steel, iron, textiles, beer and glass.

Oaxaca is a mountainous state in southern Mexico along the Pacific Ocean. Its capital city is Oaxaca, founded in 1522 as an Aztec garrison. Most of its population is Mixtec and Zapotec Indians. The economy is based in agriculture where the most important crops are: corn, wheat and rice. Tourism is also a very important income source in Oaxaca.

Puebla is mostly a mountainous state in eastern central Mexico. Its capital city is the famous city of Puebla de los Angeles where one can find more than 300 Catholic churches. Puebla is the fifth largest city in the country. Agricultural crops include corn, wheat and alfalfa. Among its industries, textiles and ceramics are important. Ceramics are famous for its blue Talavera tiles. The biggest and most important industry is an automobile assembly plant, Volkswagen, which generates many jobs and other manufacturing activities within the automobile industry.

Queretaro is located in the central plain of the country. Its capital city is Queretaro. Queretaro became an independent state in 1824. Among their most important economic activities are: livestock raising, agriculture and mining. The most important highway that connects the Federal District with the northern states of the country passes through Queretaro. Many industrial activities related to food processing can be found in its capital city.

Quintana Roo is a southeastern state of Mexico. It is one of the most recently incorporated states, this took place in 1974. Its capital city is Chetumal, characterized by having a duty free zone. The most important activity is tourism. This activity takes place in three islands off the coast, the islands are: Cancun, Cozumel and Isla Mujeres. These islands are located in what is known as the Mexican Caribbean. The state is inhabited by scattered Mayan Communities. Fishing is also an important activity as is agriculture. Its most important agricultural products are: henequen, chicle and cotton.

San Luis Potosi is an inland mountainous state in northeastern Mexico. Its capital city is San Luis Potosi. The state has important mines of silver, gold, copper, lead and zinc, among others. Agricultural activities are also important as are some industrial activities related to metal-processing plants, textile and flour mills.

Sinaloa is a northwestern state of Mexico. Its capital city is Cualiacan. Sinaloa became a state in 1830. Among its economic activities, tourism and agriculture are the most important. Its main port is Mazatlan which is also a tourist attraction. Agriculture is its most important activity, where sugarcane, cotton,citrus fruits, rice and tomatoes are the most important products.

Sonora is a mountainous state in northeastern Mexico; it borders the United States of America. It is the second largest state in area of Mexico, but at the same time it is the least densely populated. Its capital city is Hermosillo, which is a major processing center for the surrounding mining and agricultural area. Sonora has been important for its mining since the colonial times. The most important minerals are: gold, silver and lead. Crops include wheat, corn, cotton and rice. Livestock raising is also a very important activity.

Tabasco is a southeastern state of Mexico. Its capital city is Villahermosa. The most important activity is related to petroleum; Pemex, the biggest enterprise in the country which is owned by the government, has one of its most important refining and petrochemical complex in Villahermosa. Another important activity is agriculture, bananas and cacao being the most important products. Livestock raising is also a very important activity. Because Villahermosa is closer to the archeological site of Palenque which is in the state of Chiapas, and because of the different archeological sites related to the Olmeca culture, Tabasco attracts many tourists.

Tamaulipas is a mountainous state in northeastern Mexico. Ciudad Victoria is its capital city. Tamaulipas most important resource, petroleum, was discovered in the early 1900's. Tamaulipas is also important for the port of Tampico through which many of Mexican imports and exports pass. Agriculture is another very important economic activity; the predominant crops are cotton, cereals, sugarcane, tobacco and coffee. Another important activity is livestock raising.

Tlaxcala is a state located in central Mexico. Tlaxcala is the smallest and most densely populated state in Mexico. Its capital city is Tlaxcala which means "rocky place" in Nahuatl, an Indian language. Tlaxcala's most important economic activity is agricultural activities. Industrial activity in the state is related to textiles.

Veracruz is an eastern central state of Mexico, located along the Gulf of Mexico. Veracruz became a state in 1824. Jalapa is its capital city. Although its capital is Jalapa, Veracruz, a port city, is the largest urban center. Veracruz is the most important general cargo port; it handles petroleum, fruits, molasses and rum.

Veracruz's primary economic activity is agriculture. The most important agricultural products are: coffee, sugarcane, bananas, and tobacco. Livestock raising is concentrated in the lowlands. Tourism is also an important activity for the state, in which the major percentage of tourists are Mexican.

Yucatan is a southeastern state of Mexico; it occupies the north central Yucatan Peninsula. Yucatan is located in the Gulf of Mexico. Its capital city is Merida. The majority of its inhabitants are Mayan Indians. Yucatan is primarily an agricultural state where henequen is the main crop, henequen is a fiber used in making rope. Other crops are: corn, rice, sugarcane and beans. Petroleum is extracted from its coast. Another important economic activity is fishing. Tourism is also a very important economic activity that generates income for the state; archeological sites as Chichen-Itza and Uxmal are the most renowned locally and internationally.

Zacatecas is an inland state in northern central Mexico. Its capital city is Zacatecas. Its economic activity is centered in mining. Its economy is heavily dependent on gold, silver and other minerals. Other important economic activities are livestock raising and agriculture.

1.2 Efficiency

There are several criteria frequently used when the performance of a service or manufacturing organization is measured. Among the most commonly used are the following: effectiveness, efficiency, flexibility, innovation, and quality. In general it is not just measuring one dimension that gives you the broadest perspective of the performance of the organization, but rather, a variety of them.

The measure of performance for a country, from an economic point of view, could be its welfare, welfare being defined as the status of living of the inhabitants of the country, measured through the gross domestic product per capita. The United Nations (UN), a few years ago, through the United Nations Development Program (UNDP) defined an index called Human Development Index (HDI) which tries to define the level of development of countries. The index takes into account four variables related to welfare. These are: the level of longevity, literacy rate, years of education of the inhabitants, and per capita income. One of the next Chapters is devoted entirely to a detailed explanation if the HDI.

All countries, try to pursue or maintain a status of a developed country; it does not matter if the definition changes or if new variables are taken into account or incorporated to define development. Mexico, as other developing countries is no exception to the rule and pursues a developed status. Finally, what the government of the country wishes for its inhabitants is to obtain the status of a developed country because it reflects more benefits for all. Getting to the desired goal, development, would normally be measured through some index which reflects the performance. The way in which the performance is obtained, let us call it efficiency, is of extreme importance because it can define to the decision maker the amount of effort and resources that was devoted to obtaining a certain level of performance.

Measuring regional efficiency in a country is very important, because it helps the government decision makers understand and evaluate the effort and resources devoted to obtaining the development status that each country seeks. Decision makers can evaluate if too many resources were allocated for the performance obtained and can even learn from units which with less resources, performance is as good or better than the unit with which it is compared. With these ideas in mind, government or decision makers in general can be helped to better allocate resources to the different entities that demand or compete for them, resources that are inherentally scarce.

As mentioned before, there is a chapter devoted to explain the HDI proposed by UNDP which general terms calculates the degree of the development of the human resources of a country. This index is calculated using the following variables: life expectancy, number of school years, adult literacy and income which combined, give a privation index that serves to calculate the HDI.

The approach that will be followed is the technique called Data Envelopment Analysis (DEA). It calculates an efficiency measure that uses the same variables as in the HDI plus other variables that gives a broader perspective of how states in a country work to become developed. The measure of efficiency that will be obtained using DEA should be better than the HDI because projections for improvement for each state, or decision making unit (DMU), can be calculated based on other state's or DMU's, which are referent to the DMU being studied. The most important fact is that the referent DMU's reveal the best practices. This is, no assumptions have to be addressed. Another fact that explains why the DEA approach is better than the HDI is that no inferences or abstract statistical models have to be specified to calculate the efficiency or index, which in the case of HDI, it is previously defined.

The purpose of this dissertation is to measure the efficiency of the different states of Mexico for the year of 1990. A first step in pursuing the goal of this dissertation is to define and discuss the concept of efficiency according to different authors and different points of view.

Efficiency is concerned with doing things right, that is, making the right decisions to improve operations, reduce costs and increase output and quality (Evans et al., 1987).

It is also defined as making the best use of resources by minimizing waste. It means doing the job right; it is also the second criterion of job performance for the operations manager (Fogarty, Hoffmann and Stonebraker, 1989). One definition of efficiency found in Management is that efficiency is the ability to make the best use of available resources in the process of achieving goals (Bartol and Martin, 1994). Efficiency is the ratio of actual output attained to standard output expected (Sumanth, 1985). In Economy we find that efficiency is defined as the absence of waste or utilization of the economic resources that reports the maximum level of satisfaction possible with the given factors and technology (Samuelson & Nordhaus, 1993).

The DEA technique, which is based on the concept of efficiency, is extensively used in natural sciences and engineering. From the engineering point of view, efficiency is defined as the ratio of the amount of work performed by a process to the amount of energy or other inputs consumed by the process. Taking into consideration the law of conservation of energy, efficiencies ratios are less or equal to one. The engineering efficiency concept is not inmmediately applicable to the economic efficiency point of view (Yue, 1992). As Yue states, in economy output is expected to exceed inputs due to what is named the value added. This is the synergy concept in which output should be greater than the summation of its inputs of components. Charnes, Cooper and Rhodes overcame the economic value added difficulty, expected output exceeds inputs, and developed the technique known as DEA which is similar to the efficiency engineering standard. This is the law of conservation of energy, efficiencies ratios are less or equal to the unity. Energy would be transformed or lost, not gained in the process of transformation.

1.3 Its Importance

In general, in business, natural sciences, or engineering, the concept of efficiency is very important because it determines the level of inputs necessary to achieve an expected goal or level of output. Knowing what are the pursued goals, managers or decision makers must make the best effort to use the inputs with which they are endowed to achieve the goals. When the results are measured and compared with the goal, the industry standards or the competitor's results, the level of efficiency plays an important role to determine competitiveness and success.

It is important to mention that the primary goal of any organization is to be effective, do the right things, and just as important is to be efficient, that is, doing the right things right.

1.4 Relation with Productivity

The words efficiency and productivity are two of the most used in the Business and Management environment. As was mentioned before in this chapter, efficiency deals with the proper utilization of the resources in order to achieve the goal that is pursued. The proper utilization of the resources can do more with the inputs or resources that are invested or it can obtain better or increased outputs with the same inputs.

Productivity, on the other hand, is a word that has been in use for over two centuries. It was mentioned in an article written by the economist Quesnay in the year of 1766 (Sumanth 1985). A formal definition of productivity was offered in the 1950's by the Organization for European Economic Cooperation: "Productivity is the quotient obtained by dividing output by one of the factors of production. In this way it is possible to speak of the productivity of capital, investment, or raw materials according to whether output is being considered in relation to capital, investment or raw materials, etc". In general it can be found in the literature that productivity is expressed as a ratio, and this ratio is always the relation of output to input. Confusion between the terms productivity and efficiency quite often appears; productivity is concerned with the efficient utilization of resources in producing goods and/or services, while efficiency, as stated before, is the ratio of actual output attained to standard output expected (Sumanth, 1985). As it can be seen from the efficiency definition, expected standard output takes the place of inputs in the ratio to determine productivity, according to Sumanth.

Productivity can be defined in three different forms: a) partial productivity, b) total-factor productivity and c) total productivity. Partial productivity is the ratio of output to just one class of input. Total-factor productivity is the ratio of net output to the sum of associated labor and capital inputs, by net output is it meant the total output minus intermediate goods and services purchased. Total productivity stands for the ratio of total output to the sum of all input factors. It is important to mention that both input(s) and output are expressed in constant monetary currency and in some cases output has been expressed in physical units.

Of the three types of productivity mentioned above, total productivity is the one that by definition is closer to the Data Envelopment Analysis Technique. The total productivity model establishes the following equation:

$$\text{Total productivity} = \frac{\text{total measurable output}}{\text{total measurable input}}$$

where

$$\textit{Total measurable output} = vfup + vpup + voop$$

where

$vfup$ = value of finished units produced

$vpup$ = value of partial units produced

$voop$ = value of other byproducts generated.

and

$$\textit{Total measurable input} = vh + vm + vc + e + oe$$

where

vh = value of human inputs used

vm = value of material inputs used

vc = value of capital inputs used

e = value of energy inputs used

oe = value of other expense inputs used

In the broadest sense and for the purpose of this dissertation, productivity and efficiency are considered two different names to express the same measurement. Specifically, the concept of efficiency in the context of Data Envelopment Analysis has the great advantage that inputs and outputs do not have to be expressed in monetary currency terms but in the units in which they are naturally measured or expressed.

1.5 Motivation

The main objective of this research is to determine from an efficiency point of view the Mexican states that have a better capacity of transforming their inputs into some outputs that are related to the welfare of their inhabitants. Also, an important element in the motivation is pursuing the efficiency measurement through DEA, in contrast to Multiple Criteria Decision Analysis (MCDA) where the researcher or manager determines the weight to be assigned to the different elements that compose the model; this situation that permits that the researcher or manager to determine the results because of the inherent bias that any individual can perform in his estimations. Both DEA and MCDA are considered operations research techniques.

The determination of the efficiency measurement is combined with the Human Development Index (HDI) developed in the United Nations, by the United Nations Development Program (UNDP) which establishes that the development of a nation, its welfare, is achieved through health, education and income.

1.6 Statement of the Research

This research addresses the problem of determining the efficiency, from the DEA point of view,for the level of Human Development Index achieved by the thirty one Mexican states and the Federal District, for the year of 1990. A DEA model is developed for determining the efficiency accomplished by the different decision making units, the states. Specifically, the research concentrates on developing a model that serves to:

1. Determine the relative HDI for each Mexican state.

2. Determine an optimal stratification of Mexican states based on educational and socioeconomical variables .

3. Determine the capacity of transforming specific inputs to outputs, which will be the measure of efficiency. This measure will be explained in greater detail in the chapter that explains the DEA technique.

4. Determine the maximal proportional increase in output levels and decrease in input levels which inefficient states should present if they were to become efficient.

5. Determine the referent set or facet for each inefficient state, which is represented by a subset of the efficient states.

6. Determine the substitution rates between inputs, between outputs when it is possible, also between inputs and outputs and recommend the best social expenditure policy for the states.

1.7 Organization of the Dissertation

The rest of this dissertation is organized in the following way:

-Chapter 2 describes the technique to obtain the Human Development Index (HDI), its principles, its characteristics and limitations are explained.

-Chapter 3 describes the DEA technique, its principles, characteristics, different versions, and limitations.

-Chapter 4 presents the variables used to compute the HDI the sources from which the variables were obtained, the results for each state and a discussion of those results.

-Chapter 5 presents an optimum stratification technique used to classify Mexican states and the comparisons of the result of this technique with the results obtained in the HDI.

-Chapter 6 presents the results of the application of the DEA technique. It starts by explaining the inputs and outputs variables used, the sources from which they were obtained, and finally a discussion of the results.

-Chapter 7 presents contributions and conclusions of this dissertation and establishes future research related with this project, as well as the limitations that it has.

2. The UN Methodology

2.1 Introduction

Human development has become the central point in relation to the debate of international development. The idea of human development has been incorporated to the strategies of development of several countries all over the world e.g. Bangladesh,Colombia, Ghana and Pakistan. It has become a central focus, in the 90's, of the international strategy adopted by the United Nations for development (Desarrollo Humano, 1992).

Since 1990, the United Nations Development Program (UNDP) introduced a new concept related to development. The main purpose was to measure the level of development of the countries through an index called the Human Development Index (HDI). This index is composed of several variables related to longevity, education and income. With the calculation of the index, a number between 0 and 1 is obtained, expressing the level of development of all the inhabitants of the country. The idea of development, studied in a broad perspective can mislead the conclusions of the development when they are associated with the individuals that form part of the society of a specific country. The measures obtained when the development study is expressed in a macro sense can lead one to think that all the individuals of a society have the same status, a situation that obviously doesn't hold because the index is a central tendency measure of the whole population. All citizens of a given developed country do not have the same level of opportunities, health, jobs, income etc.; the same is true for developing countries. In this sense the HDI is a new perspective to evaluate the level of development of countries based on what the average population of that country can achieve through the opportunities that the country offers.

HDI is calculated taking the following elements into account: national income, years of life expectancy and education. The developers of the index considered that if an individual is healthy and has education, he can generate income; if a country can offer these opportunities to its active economic population then such a country would have a high HDI.

2.2 Human Development Index Model

The index developed by UNDP includes three essential components: longevity, knowledge and income. These three essential components are combined in a three step process to calculate an average privation index which serves to calculate the HDI (Informe Sobre Desarrollo Humano, 1993).

Longevity is measured through the average years of life expectancy at birth. This variable, as the rest of the variables that are calculated to obtain the HDI, is calculated through what is known as a relative distance. The relative distance is the relation between the difference of the average life expectancy of the country which is being measured minus the life expectancy of the country with the minimum life expectancy of the set of countries that are part of the sample, and the difference between the maximum life expectancy of the set of countries minus the minimum life expectancy of the set. This relation is set up to calculate what is known as the longevity privation index. The algebraic relation is the following:

$$H_{ij} = \frac{\max\limits_{k} X_{ik} \quad X_{ij}}{\max\limits_{k} X_{ik} - \min\limits_{k} X_{ik}}$$

where:

H_{ij} = Privation index, where i represents the dimension being measured (life expectancy, educational achievement or GDP) and j the country for which the index is being calculated.

X_{ij} = Value of the dimension i of the country j

X_{ik} = Minimum or Maximum value of the dimension i of the countries that are part of the sample

Educational achievement is measured through two variables that are related to education. These variables are: percentage of adult literacy and the median schooling years of the adult population. Citizens of 25 years or older are considered the adult population. The two variables which are multiplied by a weight of a_1 and a_2 respectively produce the variable named educational achievement.

$$E_i = a_1 D_i + a_2 F_i$$

where:

E_i = educational achievement of country i

D_i = percentage of adult literacy of country i

F_i = median of school years of adult population of country i

a_1 = weight of 2/3

a_2 = weight of 1/3

The first HDI report published by the UN in 1990 just took into account adult literacy as the variable that captured the education level of a country. That variable by itself was not a good discriminant between industrialized countries, so the next step was to incorporate in their future reports another variable that helped to discriminate better.

Median of adult school years was the new variable that, combined with adult literacy, helped in the discrimination process. Educational achievement then was the result of the weighted combination of adult literacy and median of adult school years; the weight that gave the best discrimination results was 2/3 for adult literacy and 1/3 for adults school years.

For each country educational achievement is calculated and with that variable a privation index of education achievement H_{ij} is determined..

The third element taken into consideration for calculating the HDI is a variable related to income. The variable through which the income is measured or calculated is the Gross Domestic Product (GDP) per capita. The Atkinson equation (Informe Sobre Desarrollo Humano, 1993), which has as its dependent variable *W(y)*, is used to obtain the profit or welfare derived of *y*, income or GDP per capita; it is calculated based on the premise of diminishing returns. The Atkinson equation for each segment is:

$$W(y)=\frac{1}{1-\varepsilon}y^{1-\varepsilon}$$

In general

$$ay^* \leq y \leq (a+1)y^* \Rightarrow \varepsilon = \frac{a}{a+1}$$

ε = elasticity of marginal utility of the income related to income. $\varepsilon = a/(a+1)$, where a represents the number of times that the income surpasses the poverty threshold. In the case that the income is less than the poverty threshold, $\varepsilon = 0$. In the case that the income is greater than the poverty threshold but less than two times the poverty threshold a = 1 which makes $\varepsilon = 1/2$. In general, the value of ε depends on the range in which it is contained:

$$0 \leq y \leq y^* \Rightarrow \varepsilon = 0$$

$$y^* \leq y \leq 2y^* \Rightarrow \varepsilon = \frac{1}{2}$$

$$2y^* \leq y \leq 3y^* \Rightarrow \varepsilon = \frac{2}{3}$$

$$3y^* \leq y \leq 4y^* \Rightarrow \varepsilon = \frac{3}{4}$$

So, to obtain the welfare derived of a specific income, a summation has to be done

$$w(y) = \sum_i w(y_i)$$

as an example we can see the form of Atkinson's equation for different levels of income:

$$\text{I: } W(y) = y \qquad y \leq y^*$$

$$\text{II: } W(y) = y^* + 2(y - y^*)^{1/2} \qquad y^* \leq y \leq 2y^*$$

$$\text{III: } W(y) = y^* + 2\sqrt{y^*} + 3(y - 2y^*)^{1/3} \qquad 2y^* \leq y \leq 3y^*$$

where:

W(y) = profit or welfare obtained with the y income

y = income

y^* = poverty threshold

The welfare curve obtained with the Atkinson equation is segmented, presenting diminishing returns. When calculating the welfare derived of a level y of income, the first thing to be done is to compare income (y) with the poverty threshold to check if the income is less than the poverty threshold, if its greater than the threshold, or if its less than two times the threshold and so on. Depending on the results of the income compared to the poverty threshold, it determines the limits of the summation. The welfare is the result of the calculation of the previous segments plus the diminishing increase of the segment within which the income falls.

Once the *W(y)* is calculated for each country the next step is to calculate the GDP privation index through equation (1).

With the three privation indexes calculated before, for longevity, education and income respectively, an average privation index is calculated. The equation for the average privation index becomes:

$$H_j = \frac{1}{3}\sum H_{ij}$$

where:

H_j = average privation index of country j

Finally, the HDI_j is calculated as the difference of 1 minus the average privation index:

$$HDI_j = 1 - H_j$$

The HDI_j is visualized as a welfare index. The welfare index HDI should tend towards 1 in the countries with the highest relative level of welfare or alternatively with the lowest privation indexes.

3. The DEA Technique

3.1 Background

The Data Envelopment Analysis Technique, better known as DEA, began in 1978 with Edwardo L. Rhodes Ph. D. dissertation at the Public Affairs School at Carnegie Melon University, under the supervision of W. W. Cooper (Charnes, Cooper, Lewin & Seiford, 1994). The main purpose of the dissertation was to evaluate the Program Follow Through, an educational program for disadvantaged students. The purpose of this analysis was to evaluate and compare the performance of school districts which were participants of the Program Follow Through with school districts that were not participating in the program. They were trying to measure what is known in economics as the technical efficiency but without information about the prices of the factors. The dissertation gave birth to the first paper in which DEA as a technique was utilized by its authors (Charnes, Cooper & Rhodes). Previous to the publication of this first paper, efficiency was measured considering only one input and one output at a time. The authors worked to calculate an efficiency index where multiple inputs and outputs were considered. This resulted in the CCR ratio model.

The CCR model used Linear Programming to generalize Farrell's 1957 previous work related to Technical Efficiency in which one input and one output were considered. Charnes, Cooper and Rhodes also considered one virtual input and one virtual output which were formed of multiple inputs and multiple outputs respectively. DEA was born as a Management Science technique for the analysis of technical efficiency of decision making units (DMU) of the public sector. DEA as a technique has become an alternative and complementary approach to central tendencies measures. The units, better known as Decision Making Units (DMU), can be units within an organization, e.g. departments or organizations themselves, e.g. companies, states or even countries.

In contrast to the regression approach, in which the objective is to find a function that represents all the data by determining a function that minimizes the distance between the observation and the line or plane that is represented by the regression equation, the DEA's objective is to determine a discrete piecewise frontier, which is determined by the set of Pareto-efficient DMU's. By Pareto Efficiency we mean, in the output orientation, a DMU is efficient if it is not possible to increase any output without increasing any input and without not decreasing any output; in the input orientation, a DMU is efficient if it is not possible to decrease any input without not augmenting any other input and without not decreasing any output (Charnes, Cooper & Rhodes, 1981).

A very important issue that makes DEA a very complete technique, is that problems or decision making situations that are of a multi-criteria nature are subject to be analyzed effectively using this technique. The DEA can be considered a fair technique in the sense that the decision maker, who selects a course of action, does not establish apriori a weight for each of the elements to be considered in the evaluation of each unit. The technique searches for the weights that are more convenient for the DMU which is being evaluated but without the objective of worsening the other DMU's. Reinforcing the idea of fairness of the DEA technique, it seems that it assumes that resources across the diferent units are uniform in their characteristics.

3.2 The Model

DEA's original statement is a Fractional Programming Model that is converted to an equivalent Linear Programming model (LP); it will derive an efficiency score between

0 and 1 for each of the DMU's. The DEA technique incorporates the ability to compare the effciency of multiple service units which uses multiple inputs to produce multiple outputs; in simple words a multicriteria situation. Expressed in other words, the DEA technique chooses a scoring function for each decision making unit. Based on linear programming the DEA maximizes the score of each decision making unit, with the restriction that the scores of each of the units, including the one that is being maximized, must be equal to or less than 1.

DEA seeks to derive an efficiency measure for a DMU of the following form:

$$\frac{a\,O_1 + bO_2}{cI}$$

This equation derives an efficiency measure for a DMU with two inputs and one output where:

I = input

O_i = output i

a = coefficient of output 1

b = coefficient of output 2

c = coefficient of input

and where we are allowed to choose the values of a, b and c to give the maximun possible value of the ratio for the DMU. This is subject to the value of the ratio of efficiency for any other unit to which it is compare constrained to a maximum of 1.

A general model can be derived for the cases where multiple inputs and outputs are elements of the ratio of efficiency that is to be evaluated.

In algebraic terms:

Objective Function:

Maximize

$$E_o = \frac{\sum_{j=1}^{s} w_j y_{jo}}{\sum_{i}^{r} v_i x_{io}}$$

Subject to:

$$\frac{\sum_{j=1}^{s} w_j y_{jm}}{\sum_{i=1}^{r} v_i x_{im}} \leq 1; m = 1, \ldots, n$$

$$w_j \geq 0; j = 1, \ldots, s$$

$$vi \geq 0; i = 1, \ldots, r$$

Outputs are represented by y_j, where $_j$ goes from 1,....,s;

Inputs are represented by xi where $_i$ goes from 1,...., r;

w_j represents the weight that the model calculates for each output and

v_i represents the weight that the model calculates for each input.

This is a model which is recognized as a fractional linear programming model CCR (1978) which can be reformulated to become a standard linear programming model. The objective function, E_O, represents a ratio for the unit $_O$ that is been evaluated. The ratio is constrained so that it can only take values smaller than or equal to 1. In the case that the ratio is 1, the unit represents an efficient unit; in other cases it will represent an inefficient unit. When the unit being evaluated is inefficient, a subset of the other units whose value of the ratio is 1 represent the reference set for the inefficient DMU that is being evaluated. On behalf of using the DEA technique, a reformulated linear programming model should be solved for each DMU that is being evaluated. The transformed linear model is similar for each DMU except for the objective function and a restriction related to the DMU that is being evaluated. This original fractional model is subject to a set of constraints.

The model calculates those weights for the inputs and outputs trying to make the efficiency index in the objective function as close to 1 as possible subject to the above constraints.

The transformed fractional model in its primal linear programming formulation becomes:

Objective Function:

Maximize

$$E_o = \sum_{j=1}^{s} w_j y_{jo}$$

Subject to:

$$\sum_{i=1}^{r} v_i x_{im} - \sum_{j=1}^{s} w_j y_{jm} \geq 0; m=1,\ldots,n$$

$$\sum_{i=1}^{r} v_i x_{io} = 1$$

$$w_j \geq 0; j=1,\ldots,s$$

$$v_i \geq 0; i=1,\ldots,r$$

The dual formulation of the above model is:

Objective Function:

Minimize

$$f_O$$

Subject to:

$$-\sum_{m=1}^{n} L_m x_{im} + f_o x_{io} \geq 0, i=1, \ldots, r$$

$$\sum_{m=1}^{n} L_m y_{jm} \geq y_{jo}, j=1 \ldots, s$$

This dual model can be interpreted as finding the minimum efficiency f_O for the unit $_O$. such that, each input, a weighted combination of all units can be found in which the combined input does not exceed the proportion of the objective function of the input of unit $_O$, and for each output the weighted combination of output is at least as great as that of the unit been evaluated, $_O$. The set of DMU's with positive weights ($L_m > 0$) in the evaluation of unit $_o$ are its reference set, if $L_o = 1$ means that unit $_o$ is efficient and this unit does not have any other reference set other than itself.

3.3 The Different Types of Efficiency

DEA, the technique just explained before has the possibility of measuring three empirically-derived concepts related to production economics or efficiencies.

There are three types of efficiencies that can be measured using DEA, these are Technical, Allocative and Overall efficiencies (Sueyoshi, 1992). The technical efficiency (TE) is a measure of the proportion of the inputs used to produce the outputs. Allocative efficiency (AE) can be measured if information related to input prices can be obtained. So AE measures the proportion of actually incurred costs of inputs to minimum possible costs of inputs. Overall efficiency (OE) is the measure of the combination of TE and AE.

$$OE = TE * AE$$

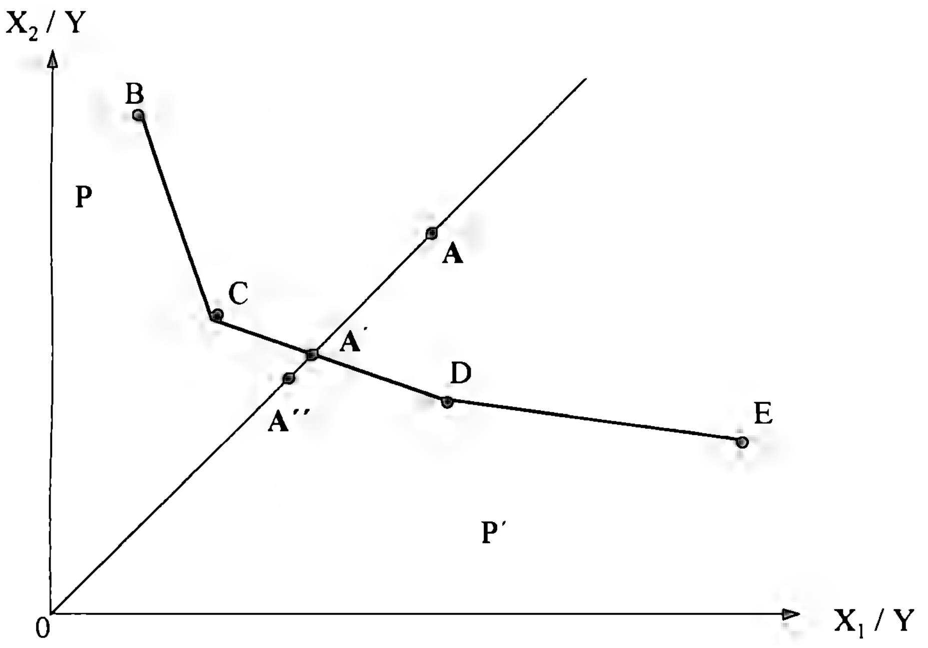

Figure 3-1 Graphic describing measurement of efficiencies.

The figure above represents the efficiency frontier for a situation where five DMU's exist: A, B, C, D and E. The efficiency measure is obtained through the proportion of a single output y to two inputs x_1 and x_2. In the horizontal coordinate the x_1 per unit output level is expressed and in the vertical coordinate the x_2 per unit output level is expressed. Points B, C, D and E represent a TE = 1, because they are part of the frontier. Point A represents an inefficient DMU because it is enveloped by the frontier.

The TE of point A is measured as:

$$TE=\frac{0A'}{0A}$$

The AE is calculated taking into account the slope of the line PP′ which represents information of the inputs prices. The AE of point A is measured as:

$$AE=\frac{0A''}{0A'}$$

Finally, OE in the figure can be measured as :

$$0E=\frac{0A''}{0A}$$

3.4 Characteristics

In general, as it was mentioned before, DEA is a non-parametric technique which calculates the efficiency of the DMU′s under evaluation. One of its most important characteristics is that efficiency, TE, AE or OE, can be calculated for cases with multiples inputs and outputs; the efficiency obtained has values between 0 and 1. Another very important characteristic is that inputs or outputs with different dimensions can be added because of the coefficients that multiply each of them. These coefficients can be normalized to sum one and their dimensions are such that the product of the coefficient by its input or output makes the result of the product one without dimension. It can be stated that DEA calculates an index of efficiency and it can be interpreted as the capacity of a DMU to transform inputs into outputs.

3.5 DEA Models

Different DEA models have been developed, among them we find the Additive model (Charnes et. al., 1985),), the Multiplicative model (Charnes et. al., 1982, 1983), the Banker, Charnes and Cooper (BCC) model (1984) and the Charnes, Cooper and Rhodes (CCR) ratio model (1978).

3.5.1 Additive Model

The additive model relates DEA to the early Charnes Cooper (1959) inefficiency analysis and relates the efficiency results to the economic concept of Pareto Optimality, which considering an output orientation means that a "DMU is not efficient if it is possible to augment any output without increasing any input and without decreasing any other output" (Charnes, Cooper and Rhodes, 1981). The optimal value of the linear program yields an efficiency value that measures the distance that the particular DMU that is being rated lies from the frontier.

$$
\begin{aligned}
&\min_{l,s^+,s^-} Z_o = -\vec{1}s^+ - \vec{1}s^- \\
&s.t. \\
&Yl - s^+ = Y_o \\
&-Xl - s^- = -X_o \\
&\vec{1}l = 1 \\
&l, s^+, s^- \geq 0
\end{aligned}
$$

The above linear program is the primal of the Additive Model. In this model a DMU is efficient if $Z_o = 0$; it is inefficient if any component of the slack variables (s^+,s^-), is not zero. The inefficiency in the outputs and inputs are determined by the values of the slacks that are not zero. This serves to identify the sources and amounts of the inefficiencies. This model presents variable returns to scale due to the constraint of lambdas (λ) which sums to 1. The geometry of this model presents a piecewise envelopment. The following figure represents a graph of the envelopment surface of the additive model.

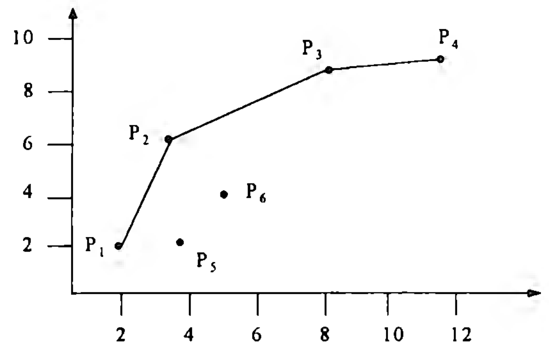

Figure 3-2 Envelopment surface for the additive model.

The envelopment surface for the additive model and the BCC input model are the same; the difference consists in the way an inefficient DMU is projected to the envelopment surface. In the case of the additive model, it selects on the envelopment surface the distance in the northwesterly direction, for the case of the figure is P_2.

3.5.2 Multiplicative Model

The multiplicative model results of the application of logarithms to the original values. Its primal linear programming model is the following:

$$
\begin{aligned}
&\min_{\lambda, s^+, s^-} Z_o = -\vec{1}s^+ - \vec{1}s^- \\
&s.t. \\
&Lo\vec{g}(Y)\lambda - s^+ = Lo\vec{g}(Y_o) \\
&Lo\vec{g}(X)\lambda - s^- = Lo\vec{g}(X_o) \\
&\vec{1}\lambda = 1 \\
&\lambda, s^+, s^- \geq 0
\end{aligned}
$$

The geometry of the model results in a piecewise log-linear or a Cobb-Douglas envelopment. All the interpretations and comments of the additive model hold for the multiplicative model. We can find two types of multiplicative models, the invariant and the variant. The difference in the latter is that the λ constraint does not form part of the linear program, a situation that forces the envelope to pass through the origin. Below is a figure which represents the graphical form for a multiplicative model.

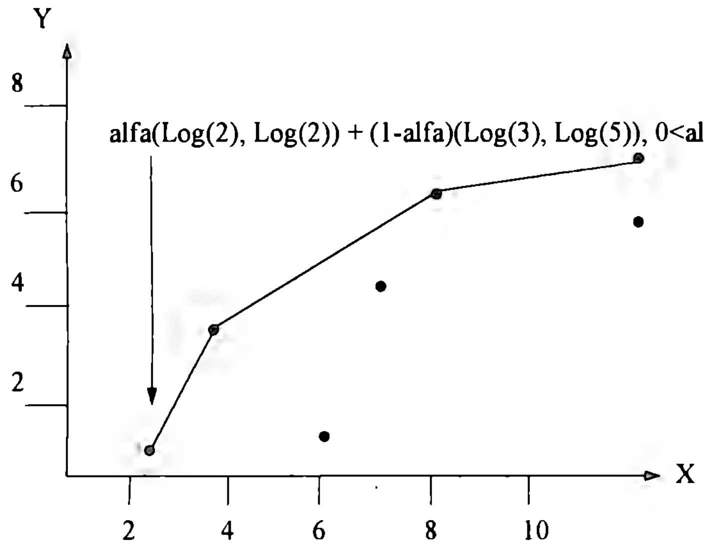

Figure 3-3 Cobb-Douglas Envelopment Surface

The formula in the figure above represents the facet of the first part of the piecewise envelopment surface as a Cobb-Douglas algebraic expression.

3.5.3 Banker, Charnes and Cooper (BCC) Model

This model has both an input and output orientation version. The former focuses on maximal movement toward the frontier through proportional reduction of inputs, in the case of the latter, the focus is maximal movement through augmentation of outputs.

3.5.3.1 **Banker, Charnes and Cooper (BCC) Input Orientation Model**

The following model represents the primal of the BCC with an input orientation. The presence of the non-Archimedian ε in the primal objective function allows the minimization over θ to preempt the optimization involving the slacks. The optimization is computed in two stages with maximal reduction achieved first through the calculation of θ, and in the second stage movement onto the efficient frontier via the slacks.

$$\min_{\theta,\lambda,s^+,s^-} Z_o = \theta - \varepsilon \cdot \vec{1} s^+ - \varepsilon \cdot \vec{1} s^-$$
$$s.t.$$
$$Y\lambda - s^+ = Y_o$$
$$\theta X_0 - X\lambda - s^- = 0$$
$$\vec{1}\lambda = 1$$
$$\lambda, s^+, s^- \geq 0$$

In this model a DMU is efficient if the following two conditions are satisfied:

a) $\theta^* = 1$

b) all slacks are zero.

The following figure presents the envelope surface in the case of a BCC input model. The points P_1, P_2, P_3 and P_4 represent efficient DMU's, which as a group conform the envelope surface. The points P_5, P_6 and P_7 represent inefficient DMU's, their position in the graph also define their DMU's reference set, e.g. P_7 reference set is P_3 and P_5 reference set is P_1 and P_2.

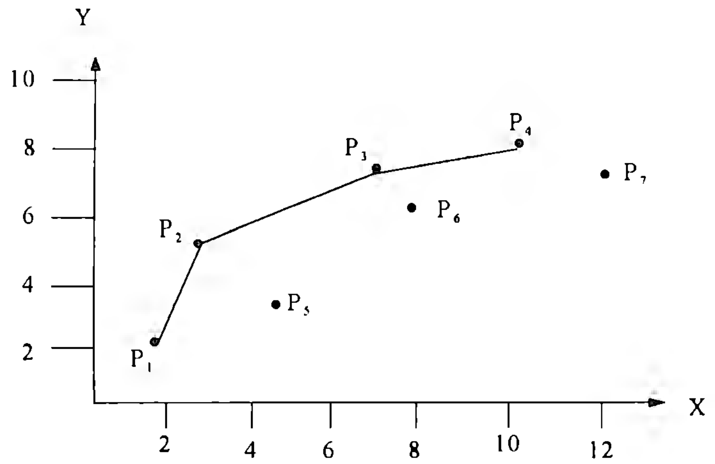

Figure 3-4 Envelopment surface for the input-oriented BCC model.

3.5.3.2 Banker, Charnes and Cooper (BCC) Output Orientation Model

The following model represents the primal of the BCC with an output orientation. The presence of the non-Archimedian ε in the primal objective function allows the maximization over ϕ to preempt the optimization involving the slacks. The optimization is computed in two stages with proportional output augmentation achieved first through the calculation of φ, and in the second stage movement onto the efficient frontier via the slacks.

$$\max_{\vartheta,\lambda,s^+,s^-} Z_o = \vartheta + \varepsilon \cdot \vec{1} s^+ + \varepsilon \cdot \vec{1} s^-$$

$$\begin{aligned}
& s.t. \\
& \vartheta Y_0 - Y\lambda - s^+ = 0 \\
& X\lambda - s^- = X_o \\
& \vec{1}\lambda = 1 \\
& \lambda, s^+, s^- \geq 0
\end{aligned}$$

In this model a DMU is efficient if the following two conditions are satisfied:

a) $\phi^* = 1$

b) all slacks are zero.

The following figure presents the envelope surface in the case of a BCC output model. The points P_1, P_2, P_3 and P_4 represent efficient DMU's, which as a group conform the envelope surface. The points P_5, P_6 and P_7 represent inefficient DMU's, their position in the graph also define their DMU's reference set, e.g. P_7 reference set is P_4 and P_5 reference set is P_2 and P_3.

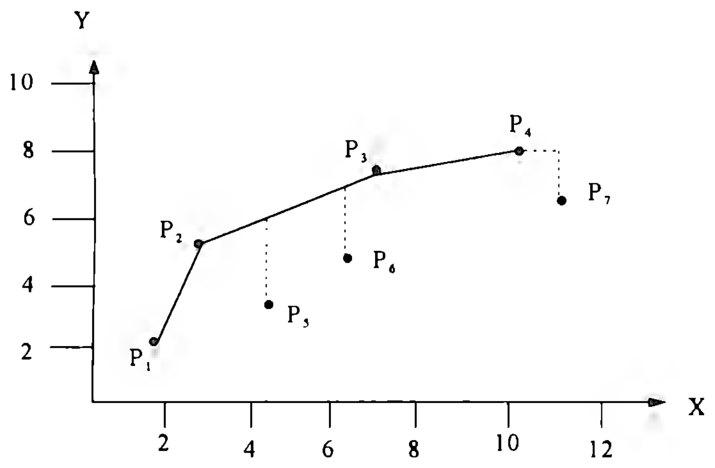

Figure 3-5 Envelopment surface for the output-oriented BCC model.

3.5.4 Charnes, Cooper and Rhodes (CCR) Model

The CCR was the first DEA model that was developed. The model admits the input and output orientation. The model assumes constant returns to scale and this is the reason why the envelope surface is represented by a 45° slope as is presented in the following figure. The primal LP for the CCR input model is the following:

$$
\min_{\theta,\lambda,s^+,s^-} Z_o = \theta - \varepsilon \cdot \vec{1} s^+ - \varepsilon \cdot \vec{1} s^-
$$

$$
\begin{aligned}
& s.t. \\
& Y\lambda - s^+ = Y_o \\
& \theta X_0 - X\lambda - s^- = 0 \\
& \lambda, s^+, s^- \geq 0
\end{aligned}
$$

An important contrast between this model and the BCC is the fact that the convexity constraint is relaxed; it is not part of the model. In this model it may be neccesary to augment some outputs and reduce some inputs to achieve efficiency.

The following figure presents the envelope surface in the case of a CCR input model.

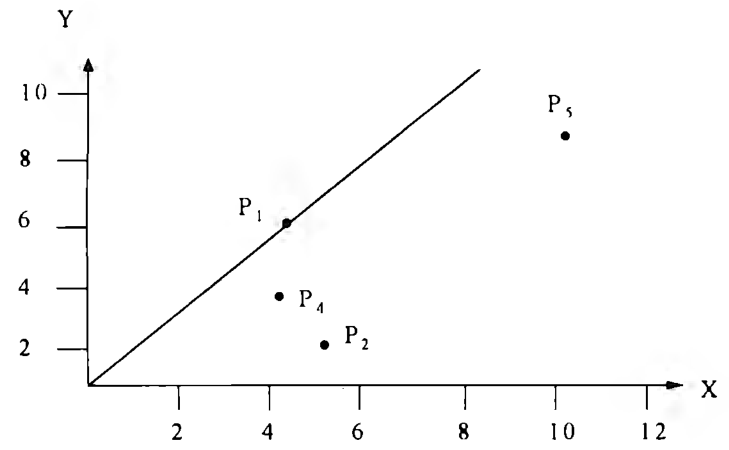

Figure 3-6 Envelopment surface for the input-oriented CCR model.

3.6 Sensitivity Analysis

In general, any LP model or derivation of it permits deeper understanding of the model when sensitivity analysis is performed. DEA models are no exception to this rule. By sensitivity analysis we mean perturbation of the original data to determine to what extent results vary. Perturbation of the original data can be made through:

a) Changing the objective function coefficient of a nonbasic variable

b) Changing the objective function coefficient of a basic variable

c) Changing the right-hand side of a constraint

d) Changing the column of a nonbasic variable

e) Adding a new variable or activity

f) Adding a new constraint.

DEA models are based on input and output variables and as we know efficiency is what is measured. Then, the main idea to achieve efficiency is decreasing as much as possible the usage of inputs and or increasing the output achievements. Sensitivity analysis permits answering questions such as: a) what is the effect on efficiency by decreasing a unit of an input? b) what is the effect on efficiency by increasing a unit of an output?. Other interesting questions would be: a) what is the magnitude of the change in efficiency if we decrease input a instead of input b? b) what is the magnitude of the change in efficiency if we increase output c instead of output d? c) what is the magnitude of the change in the efficiency of a DMU if we rather increase output c instead of decreasing input a or vice versa? To answer these questions calculation of substitution rates between inputs, between outputs and between inputs and outputs respectively is necessary. All these calculations permit one to identify the best policy or order of inputs to be decreased, outputs to be increased and the combination between them.

3.7 Literature Review

3.7.1 History

DEA, an operations research technique, had its first formal appearance in a dissertation entitled: "Data Envelopment Analysis and Approaches for Measuring the Effiicency of Decision-making Units with an application to Program follow-through in U.S. Education." The dissertation research was done by Edwardo Rhodes in 1978, when he was a Ph. D student at the school of Urban and Public Affairs in Carnegie Mellon University. Rhodes was under the supervision of Professor W. W. Cooper (Charnes, Cooper, Lewin and Seiford, 1994). So it can be said that the birth of the technique is with this dissertation.

The dissertation mentioned above gave birth to what today is known as the CCR model, named that way because of the authors of the first paper related to DEA Charnes, Cooper and Rhodes. A few years later, another DEA model appears; this one is known as the BCC, also because of its authors Banker, Charnes and Cooper. Other milestones that are important to be mentioned because of their importance are: Window Analysis, the Multiplicative and Additive models, and the incorporation of categorical and nondiscretionary variables in the model such as the stochastic and alternate frontiers.

3.7.2 Review of Research Applying DEA

This dissertation relates the DEA technique to HDI; this is by measuring the efficiency of Mexican states in achieving their specific development index. The first and most important source in the action of reviewing the DEA literature was obtaining a paper that could be as broad as possible and could give an understanding of the application of the technique. This was version 5.0 of the Bibliography of Data Envelopment

Analysis, (Seiford, 1990). In the review, just one paper was directly related with the topic covered in this dissertation, "Using DEA to Evaluate the Efficiency of Economic Performance by Chinese Cities", (Charnes, Cooper & Li, 1989).

The objective of the paper to use DEA as a tool that served to evaluate the economic performance of 28 Chinese cities that were very important for the Chinese government economic development program. It was also intended to serve as a tool for planning the way economic program should be applied. The DEA model that was used was the CCR. In the following table we present the inputs and outputs used in this paper.

Inputs	Labor	Working Fund	Investment
Outputs	Gross Industrial Output Value	Profit and Taxes	Retail Sales

Table 3-1Inputs and Outputs variables of Charnes, Cooper and Li, 1989

The variable labor is the number of staff and workers who are engaged in industries, exclusive of farm labor. The second input is working fund which is the total annual wage bill of staff and workers and the circulating capital, and the last input is investment. The first output is the gross industrial output value which is a measure of the contribution to the national economy, and it is expressed as the total monetary value of outputs. The second output variable is profit and taxes that is a measure of the net contribution to the central government; this contribution is the amount of profit and taxes generated by state-owned enterprises. Finally the third variable is retail sales which is a measure of the local consumption generated in the city (DMU) that is given by the total sales value of the products in the local market.

Among the results obtained in this study we can mention that DEA helped to identify technical inefficiencies and the waste in the activities of each city, also insights on the expansion or contraction of the economic activities of the cities so as to obtain full efficiency.

A second paper with a direct relation to the topic covered in this dissertation is one related to regional concentration and efficiency in the Mexican manufacturing sector (Bannister & Stolp, 1995). The purpose of the paper was to investigate the link between industrial location, concentration and economic efficiency in Mexican manufacturing in the year of 1985. The authors measured the technical, allocative, overall and scale efficiency of the manufacturing sector in Mexican states to observe if concentration of firms caused diseconomies of concentration. The DEA technique was used to calculate the different efficiencies of the firms in each state Also a regression model was made to relate the efficiencies with the characteristics of the manufacturing sector of the states or regions. This was a case study for Mexico. The study was made for each of the seven two digit industry classification, 1)food, feed and tobacco, 2)textiles, apparel and leather, 3) wood and wooden products, 4) paper, printing and publishing, 5) chemicals, petrochemicals, rubber and plastic, 6) nonmetallic minerals, 7) metallic products, machinery and equipment. The inputs and outputs used in this study are in the following table.

Inputs	Gross fixed assets	Full time equivalent workers	Compensation
Outputs	Value added		

Table 3-2 Inputs and Output varibles of Bannister and Stolp, 1995

Results indicate that despite evidence of increasing cost of congestion, regions that are more heavily industrialized exhibit higher levels of efficiency: Mexico, Nuevo León, Jalisco and State of Mexico. These same states present decreasing returns in their efficiencies. The most important result related to the econometric study is that there is a relation between the agglomeration economy and the regional efficiency in production.

Among the different research that has been done, one area that has been studied is the relation of DEA and other techniques such as discriminant analysis and regression analysis.

Discriminant Analysis (DA), a statistical technique, is used for two main purposes: a) describing major differences among the groups in MANOVA and b) classifying subjects into groups on the basis of a battery of measurements (Stevens, 1986). The data is assumed to be multivariate normal. DA has two attractive features: a) parsimony of description and b) clarity of interpretation; these two features can be accomplished by Linear Programming (LP) or by the statistical technique process. The application of DEA seeks to classify DMU's as efficient or inefficient, which as an analogy classifies DMU's in two groups.

DEA and DA have several differences and similarities (Retzlaff-Roberts, 1996). Among the similarities we can find that both techniques involve the calculation of a score for each unit, which is obtained through a weighted sum of factors. They also both define a hyperplane that attempts to separate into groups; both methods involve a relative comparison. Among the differences we can find that DA defines groups a priori while in DEA the groups are defined after the analysis. Another difference is the one related to misclassification; in DA it is allowed while it is not desireable and in DEA there is no misclassification. DA assumes units as homogeneous while in DEA DMU's are assumed to be non-homogeneous. Depending on the answer to the following three questions, a) Is there prior knowledge of group membership? b) Are units homogeneous? and c) Is misclassification allowed into either group?, it will lead to use one or the other of the techniques. From the combination of the answers to the three questions, 8 categories result; in the case where all three answers are yes, DA is the technique recommended and if the three answers are no, DEA is recommended.

Within those extremes one or the other of the techniques is recommended with some adjustments.

DEA has been applied extensively in the educational area. Among the work in education, the academic departments efficiency of a university was measured (Sinuany-Stern et al, 1994). It was a case study of the Ben-Gurion University in Israel. The purpose of the authors was to present new directions to further analyze Decision Making Units (academic departments) and not only consider the economic aspect to evaluate each department. Two statistical techniques were also used in the study, cluster analysis and discriminant analysis. Cluster analysis was used to divide the different departments to different sets, and discriminant analysis was used to test the match of efficiency/inefficiency division that was obtained with the DEA model. In the following table the inputs and outputs used in the study are presented.

Inputs	Operational expenditure	Faculty Salaries		
Outputs	Grant money	Number of publications	Number of graduate students	Number of credit hours given by the department

Table 3-3 Inputs and Outputs variables of Sinuany-Stern et. al 1994.

The CCR DEA model was applied. The number of academic departments used in the study was 21, which were part of three of the four schools that are part of the Ben-Gurion University, (the Medical school was not included in the study). Among the more important results, the authors found that only seven departments were efficient. When several variables are combined into one, the results change significantly; all of the departments that represent a decision unit do not offer the same service (e.g. laboratory), a situation that makes that the departments that are more alike have less differences in efficiency between them than those that do not offer those services. Another important result is that the way in which the inputs and outputs are measured

has a direct impact in the results; this is, some departments are efficient depending on the way the variables are measured. The cluster analysis did not give a good fit of the efficient and inefficient departments and in the case of discriminant analysis, which was applied to each school instead to all the university, it was found that it fit very well in the discrimination of efficient and inefficient departments within a school. Finally, the authors counted the number of optimal appearances in the solution of other departments; the measure was meaningful for ranking efficient departments.
DEA has also been used to make a clear cut in order to know the efficiency achieved because of a program and how much the administration of the program helped in achieving that efficiency (Charnes, Cooper & Rhodes, 1981). One of the first articles that was published where DEA was used was the the one entitled "Evaluating Program and Managerial Efficiency: An application of Data Envelopment Analysis Follow Through." The main purpose of the article was to measure the efficiency achieved by the implemetation of two programs (program follow through and non follow through) to disadvantaged students in the public schools in the United States.

Among all the variables that could have been used to measure inputs and outputs, the authors decided to use the following variables:

Inputs	Education level of mother	Highest occupation of a family member	Parental visit index	Parent counseling index	Number of teachers
Outputs	Total reading score	Total mathematics score	Coopersmith self-esteem inventory		

Table 3-4 Input and Output variables of Charnes, Cooper and Rhodes, 1981

The model had 49 DMU's from all over the U.S. for the case of the program follow through (PFT) and 21 DMU's for the non follow through program (NFT). Three models were run, one for the PFT and another for the NFT. After running these two models where the managerial efficiency could be verified, the inefficient DMU's in each model were forced to become efficient, with the purpose of eliminating the managerial inefficiency. After that process, the third model in which all of the DMU's were combined was run. The idea of combining all the DMU's into one single model was to verify the superiority of one of both programs PFT and NFT. The evidence was that no one of the programs was superior than the other; instead, the evidence was that in certain segments of the frontier, one program achieved levels of efficiency that the other program did not.

Up to this point all of the the literature reviewed determine efficient units and inefficient ones, using different DEA models. This means that we can observe between inefficient units a ranking of the level of inefficiency situation which does not operate in efficient units. Its very important to mention the fact that one of the weakness of DEA is that considerable number of DMU's can be charaterized as efficient, unless the sum of the number of inputs and outputs is small relative to the number of observations (Andersen &Petersen, 1993; Schefczyk, 1993). We can also observe in different studies that specialized DMU's can be rated as efficient due to a single input or outputs variable, even though the respective variable may be seen as relatively unimportant. Obtaining more information out of the efficient units is desirable e.g. getting to know the maximum proportional increase in inputs of a DMU preserving the DMU efficiency (score of 1). Ranking efficient units (Andersen & Petersen,1993) permits the decision makers to distinguish economically viable units from units that are only technically efficient. The method proposed is based in the BCC DEA model and it consists of dropping the specific efficient DMU from the reference set in order to obtain an index that can be 1 or greater and its a guide to determine the maximun proportional increase in its input to maintain its status of efficient. This modification gives a decision maker information in order to allocate more inputs to efficient DMU's which will increase their outputs, maintaining its score efficiency of one and also depending on the score it will help to decide on which efficient unit, in the presence of limited resources, should investment first be made.

In the same direction of the article explained before management seeks in general to increase profitability and/or control costs while one of both becoming or remaining technically efficient in the DEA sense (Golany, Phillips & Rousseau, 1993). Three families of models developed by the latter authors were developed for managerial situations in which they can be useful; a) effectiveness improvement in an existing DMU, b) new DMU's are to enter the organization e.g. manufacturer considering the introduction of a new brand into an existing market and c) the reallocation of scarce resources among the aggregate of all DMU's for the maximum overall organizational technical efficiency and profitability. In general, to apply this model, results from a DEA study must be available. The first model uses goal programming to maximize the profit of an existing DMU. This is achieved by the projection of the inefficient DMU to a section of the efficient frontier; achieving effectiveness does not always mean increasing revenues as sometimes the opposite is what applies. The second model determines for a specific level of inputs what is the maximum attainable level of outputs that can be obtained considering that the level obtained will be as close as possible to the efficient frontier. This is achieved through goal programming. Finally, the third model of the family seeks to reallocate scarce resources between the DMU's on that profitability is increased and technical efficiency achieved. Its a five step process of running DEA to obtain efficiency scores, compute effectiveness indices, scale the indices by their averages, compute input effectiveness indices and finally solving a linear program based on the efficiency scores to weigh the effectiveness indices which guides the new allocation of resources. The DEA additive model is used as the starting point to develop each of the three models of the family.

DEA has also been applied in the service sector, hospitals in particular among others. Technical efficiency and cost effectiveness for medical care decision making, which can be calculated using DEA, is a great area of opportunity to deal with the increasing pressure of marketplace and political forces to reduce prices (Finkler & Wirtschafter, 1993). Nine hospitals that offered the obstetric services were studied through six models whose differences consisted in the aggregation of the inputs. The idea of aggregating inputs, in an attempt to avoid one of DEA´s weakness, the larger the number of inputs and/or outputs, the harder it was for DEA to determine efficient and inefficient DMU's. In other words, the more inputs and/or outputs a model has, DMU's tend to become more efficient. After running the DEA model with different levels of aggregation the next step was to compute a slope, factor/price ratio, which is a price-efficient mix of factors when the production frontier equals that slope. In other words the authors identified the allocative or input price efficiency. So the DMU's in the production frontier whose slope coincides with the factor/price ratio is technically efficient and allocativly efficient or cost effective.

Inputs	Physician hours	Registered nursing hours	Certified nurse midwife hours	Licensed vocational nursing hours	Resident hours
Outputs	Delivery of baby	Recovery of the mother			

Table 3-5 Input and Output variables of Golany, Phillips and Rousseau, 1993

The input modification choices used in the study were: a) ignore one or more inputs, b) add together factors with similar prices, c) price weight each factor and use the factor income totals per case and d) aggregate inputs, using factor prices. The most important result of the study is that the calculation of the allocative efficiency permits that cost managers obtain information in the best way to manage the DMU's.

DEA is also considered an effective tool for benchmarking (Schefczyk, 1993).

3.8 Limitations

The DEA model, as many other models, has limitations. The most important one for any study measuring efficiency deals with the sample size or the number of DMU's and the amount of variables, inputs and outputs, that are taken into account. In reviewing the literature two rules where found. The first one states that the number of DMU's should be greater than or equal to the double of the sum of inputs plus outputs. This is based on emprical findings. The equation can be stated as:

$$n \geq 2(r+s)$$

Sinuany-Stern et. al, (1994) states that the number of inputs plus the number of outputs should be less than or equal to the number of DMU's divided by three. The equation can be stated as:

$$r+s \leq \frac{n}{3}$$

In this dissertation the latter rule is used in order to obtain more robust results.

Another important limitation is the fact that DEA is concerned with efficiency results not effectiveness results which is of the most crucial importance in any decision making situation. This does not invalidate the study but it is important to take into account in order to be careful with the interpretations of the results.

4. HUMAN DEVELOPMENT INDEX (HDI) OF MEXICAN STATES

Obtaining a high welfare level for all the inhabitants of a country is, and should be, one of the most important issues in the agenda of the any government. Governments would like to know if they are achieving this objective for their citizens, so it is important to define welfare and how to measure it. In that direction a great effort has been made by the U.N. to develop an index that measures the development of the citizens of a country (which serves to measure the level of welfare).

It is important to determine the level of development of the inhabitants of a country in order to evaluate if a country is on track in providing welfare for its inhabitants and also to know where the greatest effort should be made in order to most rapidly achieve development levels that make a country wealthier and healthier.

The purpose of this chapter is to apply the HDI to the 32 Mexican states in order to obtain a ranking of their relative welfare. Explanation of the variables used, the sources which were consulted to obtain those variables and comments related to the attainmet of those variables will be given. Also included are the results obtained in the calculation of the index and a discussion of them. All the variables used for the calculation of the index are for the year of 1990 except the Gross Domestic Product (GDP), which is from 1988. The 1988 GDP was used because it better represents the welfare of the Mexicans as of 1990 than the 1993 GDP. The regional GDP in Mexico is calculated every 5 years. It used to be calculated for the years that ended in 0 or 5, but because of some changes in the measurement policy, it was calculated for 1988 and 1993. This situation leads us to use the GDP measurement for a different year than the others.

4.1 Variables

There are four basic variables that are necessary for the development of the Mexican states HDI model: rate of literacy for inhabitants of 15 years or older, educational achievement at age 16, life expectancy at birth, and income welfare, all of which will now be explained.

The rate of adult literacy is a variable that measures the percentage of the population of 15 years and older, that know how to read and write. According to INEGI (Instituto Nacional de Estatdística Geografia e Informática) a literate person is one who is 15 years or older and declares that he knows to read and write a message. This variable is known as a stock variable because it increases or decreases very slowly over time. It was calculated as the ratio of total literate inhabitants of 15 years and older to the total population of 15 years and older. The original variable is the ratio of literate adults in the population, but we considered that inhabitants of 15 years and older is probably a good measurement of adults since at age 15, Mexicans can legally join the workforce. Therefore, it was decided to calculate the ratio for 15 years and older. Another important issue is that what is considered adults in one country may not be the same for another country.

The second variable used to calculate the HDI for Mexican states was educational achievement at age 16. The variable used in the original HDI of the UN was the median of average years of school. In the case of Mexico it was very difficult to obtain the original variable because of the way data is published. After analyzing the way in which the data was presented we decided to use a variable that was a proxy to original one. This variable was the educational achievement of Mexicans students at 16 years. This variable gives us an idea of the students that are on track at that age and have good possibilities to continue with their education. The educational achievement at 16 equals one minus the ratio of population who have not completed ninth grade to the total population of 16 and older.

The two above variables are combined to produce a new variable named educational achievement. The educational achievement is obtained through the averaged weight of adult literacy and educational achievement at 16. This is:

$$E = a_1 D + a_2 F$$

where:

E_i = education achievement of state i

D_i = percentage of adult literacy of state i

F_i = educational achievement of 16 years old inhabitants of state i

a_1 = weight of 2/3 as it was used in the UNDP study

a_2 = weight of 1/3 as it was used in the UNDP study

The third variable is life expectancy at birth. This variable which is also a stock variable, because its changes very slowly, measures how long a Mexican citizen of a specific state is expected to live. The variable in some way is an index of the standard of living and the level of health in each state; as the larger the life expectancy is, the larger the index. It can be found that there are countries in the African continent whose life expectancy at birth is less than fifty years. For developed countries we find that the life expectancy is around 78 years.

The fourth and final variable used to calculate the HDI is income welfare. This is a variable that is related to income and it is measured through the use of the GDP. Using the Atkinson equation (explained in chapter 2), an income welfare variable *W(y)* is calculated. This income welfare is a variable that presents the diminishing returns phenomena. The calculation of the welfare income depends on what is called the poverty threshold. For the case of the HDI model of Mexican states, we consider the national average minimum wage.

4.2 Sources

The first of the variables mentioned, the literacy rate of persons 15 years and older was constructed or calculated taking into account: a) the total of literate persons of 15 years or older for each state in 1990 and b) the total population of persons 15 years or older of each state for 1990. The ratio of this data gives us the rate of literate persons for each state. The data was obtained of the table 21.24 on page 546 of the 1993 anual statistics published by INEGI. This report was obtained using the 1990 national population census.

The variable education achievement for population of 16 years or olders is calculated also as a ratio. To obtain that ratio we calculated the difference of one minus the ratio of population of 16 years and older in 1990 that did not have junior high school, divided by the total population of persons 16 years and older of 1990 for each state. This data was obtained from table 21.27 on page 549 of the annual statistics published by INEGI.

Life expectancy at birth for the year of 1990 is a variable for the model that is obtained from the table 21.9 on page 609 of the 1993 annual statistics published by INEGI; according to the table, the data is obtained from the interinstitutional group for health information.

The last of the four variables used to develop the index is the income welfare. The income welfare is calculated through the Atkinson's equation. This equation takes into account the poverty threshold. In the case of the Mexican states it was calculated as the average minimum wage. This general minimum wage of 1990 was obtained from the first part of table 4.14 on page 109 of the annual statistics of the Mexican states published by INEGI. The other important variable used was the GDP per capita of the Mexican states. The calculation of this second data is the result of the ratio GDP of each state divided by its population. The GDP of 1988 was published by INEGI and the total population obtained from the XI General Population and Housing Census of 1990 on page 546, was also published by INEGI.

4.3 Results

In this section we present, in different tables, the data and the results of the calculation of the variables necessary to compute the HDI for the Mexican States.
In Table 4.1 we can find the sources to compute the variable named ratio of educational achievement for population of persons 16 years and older for the year of 1990. The original variable in the U.N. study is the median of the educational achievement of the adult population. In this dissertation, the variable that was defined as a proxy to the U.N. study is the educational achievement for the population of persons 16 years and older, which is the age at which middle school is normally completed. We think that by starting at age 16, we can have a relative variable that can measure the population of persons 16 years and older who are on track to achieve a high level of education.
The calculation of the variable is based on the total population of 16 years and older and on the population of 16 years and older that do not have middle school as minimum education. In table 4.1 we can find that states such as Oaxaca, Zacatecas and Chiapas are the states with the lowest educational achievement, and in the other

State	Total population of 16 yrs. and olders 1990	Total population of 16 yrs. And olders without Junior high school 1990	Ratio of educational achievement of 16 yrs. and olders 1990
Aguascalientes	407945	307202	0.246952408
Baja California	1021029	697636	0.316732434
Baja California Sur	190758	130157	0.317685235
Campeche	302921	233289	0.229868514
Coahuila	1193020	855682	0.282759719
Colima	250734	184774	0.263067633
Chiapas	1702410	1444972	0.151219741
Chihuahua	1491897	1138185	0.237088753
D. F.	5509469	3200559	0.419080314
Durango	760095	604021	0.205334859
Edo. de México	5784180	4184531	0.276555882
Guanajuato	2192716	1803077	0.177696975
Guerrero	1401753	1120176	0.200874904
Hidalgo	1051977	845949	0.195848388
Jalisco	3073840	2340059	0.23871802
Michoacán	1954701	1545360	0.209413614
Morelos	711775	517488	0.272961259
Nayarit	466979	352042	0.246128841
Nuevo León	1954327	1279638	0.345228306
Oaxaca	1653268	1420310	0.140907584
Puebla	2309307	1820238	0.211781716
Querétaro	581795	446662	0.232269098
Quintana Roo	276929	207910	0.249229947
San Luis Potosí	1123332	883918	0.213128443
Sinaloa	1284864	910753	0.291167781
Sonora	1111986	775421	0.302670178
Tabasco	828843	647707	0.218540785
Tamaulipas	1390081	996580	0.283077749
Tlaxcala	431338	333340	0.227195378
Veracruz	3657579	2926669	0.199834371
Yucatán	810876	629803	0.223305413
Zacatecas	704420	598469	0.150408847

Table 4-1 Ratio of educational achievement for population of 16 years and older.

extreme we can find the states of Baja California, Baja California Sur, Nuevo Leon and Distrito Federal (D.F.) with the highest levels. In the lowest extreme we find ratios of 0.14 and 0.15 and at the highest extreme of educational achievement of 16 years and older we find ratios from 0.32 up to 0.42 for D.F..

In Table 4.2 we find a set of variables that are used to compute the literacy rate of the population of persons 15 years and older for the year of 1990. The original variable used in the HDI done by the UN is the adult literacy rate, but as with the educational achievement this study considers the literacy rate for persons of 15 years and older since the legal age to start working is 14 years old (Ley Federal del Trabajo, 1994). We can find that only four states have a ratio less than 80%. Those states are: Chiapas, Oaxaca, Guerrero and Hidalgo. At the other end we can find 15 states with ratios greater than 90%, with Durango, D.F., Nuevo León and Baja California being the highest with rates of 96%, 96%, 95% and 95% respectively.

In Table 4.3 we find several variables with information obtained directly from INEGI. For example we find life expectancy for each of the states, the GDP for 1988, population for 1990 and the last variable GDP per capita calculated from the two latter variables. The table is ordered increasingly according to GDP per capita; this variable expresses the average income that each citizen in his respective states earns in one full year. Three of the considered poorest states appear in the top of the list, those states are: Oaxaca, Chiapas and Guerrero, all of them southern states. At the other end, we find Campeche as the state with the highest GDP per capita, with the oil endowment and the related industry of the state being the responsible for this situation. D.F. and Nuevo Leon are the second and third respectively, as states with high GDP's.

Oaxaca, Puebla, Hidalgo, Chiapas and Guerrero are the states with the lowest life expectancy and D.F. and Nuevo Leon the states we the highest life expectancy.

State	Total population of 15 yrs. and olders 1990	Total literate population 15 yrs. and olders 1990	Literacy rate of 15 yrs. and olders 1990
Aguascalientes	426272	395790	0.928491667
Baja California	1060070	1007665	0.950564585
Baja California Sur	198475	187036	0.942365537
Campeche	316000	266668	0.843886076
Coahuila	1243021	1173503	0.944073350
Colima	261855	237203	0.905856295
Chiapas	1779514	1238682	0.696078817
Chihuahua	1550954	1453411	0.937107741
D. F.	5697831	5462684	0.958730436
Durango	796484	740130	0.929246539
Edo. de México	6031182	5487799	0.909904394
Guanajuato	2296303	1910604	0.832034797
Guerrero	1470387	1073014	0.729749379
Hidalgo	1099122	869884	0.791435346
Jalisco	3207598	2917323	0.909503934
Michoacán	2049007	1687462	0.823551115
Morelos	742676	653431	0.879833198
Nayarit	489418	432194	0.883077451
Nuevo León	2030607	1933702	0.952277816
Oaxaca	1727382	1249297	0.723231457
Puebla	2411512	1943675	0.805998477
Querétaro	608193	513609	0.844483577
Quintana Roo	288019	251742	0.874046504
San Luis Potosí	1174199	996753	0.848879108
Sinaloa	1345071	1209781	0.899417949
Sonora	1156950	1088944	0.941219586
Tabasco	865525	754206	0.871385575
Tamaulipas	1445656	1344595	0.930093328
Tlaxcala	450368	399891	0.887920545
Veracruz	3806601	3106903	0.816188248
Yucatán	844997	709699	0.839883455
Zacatecas	739627	665787	0.900165894

Table 4-2 Population of 15 years and older literacy rate.

State	Life expectancy 1990	GDP 1988	Total Population 1990	GDP per capita
Oaxaca	62.1	6,690,090	3019560	2,215,584
Chiapas	66.4	7,592,696	3210496	2,364,960
Michoacán	68.7	9,790,101	3548199	2,759,175
Guerrero	67.4	7,352,077	2620637	2,805,454
Tlaxcala	67.6	2,230,531	761277	2,929,986
Puebla	65.9	12,121,622	4126101	2,937,791
Zacatecas	71.3	4,009,839	1276323	3,141,712
Guanajuato	69.3	12,894,944	3982593	3,237,826
Yucatán	70.7	4,551,061	1362940	3,339,150
Nayarit	71.2	2,842,496	824643	3,446,941
Hidalgo	66.0	6,646,161	1888366	3,519,530
Veracruz	69.3	22,183,413	6228239	3,561,747
San Luis Potosí	70.8	7,218,160	2003187	3,603,338
Durango	71.3	5,141,981	1349378	3,810,631
Aguascalientes	71.2	2,854,161	719659	3,965,991
Sinaloa	71.9	8,763,790	2204054	3,976,214
Morelos	70.6	4,984,828	1195059	4,171,198
Edo. de México	70.1	44,511,815	9815795	4,534,713
Jalisco	70.9	26,463,416	5564892	4,755,423
Tamaulipas	71.4	10,704,309	2249581	4,758,357
Tabasco	67.7	7,250,811	1501744	4,828,260
Colima	69.4	2,085,861	428510	4,867,707
Querétaro	68.3	5,124,193	1051235	4,874,451
Chihuahua	69.8	12,684,878	2441873	5,194,733
Quintana Roo	71.3	2,815,426	493277	5,707,596
Baja California Sur	71.4	1,821,170	317764	5,731,203
Sonora	70.2	10,746,400	1823606	5,892,940
Coahuila	71.4	11,664,132	1972340	5,913,855
Baja California	70.8	9,929,109	1660855	5,978,312
Nuevo León	74.1	24,721,058	3098736	7,977,788
D. F.	73.2	83,371,073	8235744	10,123,077
Campeche	71.8	8,689,697	535185	16,236,810

Table 4-3 Life Expectancy, GDP for 1988 and GDP per Capita

In Table 4.4 we can find the different indices that are used to compute the HDI for each of the states. In the first column we find the poverty threshold, which is based on the Atkinson equation, used to calculate the welfare based on the income. This is an equation, as stated in Chapter 2, that presents diminishing returns. In the second column we find the life expectancy privation index, in the third column we find the education privation index, in the fourth column we find the income privation index, and in the fifth column we find the average of the three previous columns. This is the average privation index for each state. In the last column we will find the HDI for each of the states. We can observe in this table that the worst indices (HDI) are for the states of Oaxaca, Chiapas and Guerrero, and the states with the highest relative HDI are: D. F., Nuevo Leon , Baja California and Baja California Sur. In the second and third columns we will find cells that have a dash. This means that those states were selected by the method as the benchmark to obtain the respective index. Table 4.4 is presented in the ascending order of its last column. This is from the state of the worst HDI to the state with the best relative HDI.

To finish this section the state of Morelos will be used as an example to demonstrate the way the different variables and indexes were computed.

The calculation of the Ratio of Educational Achievement of the population of 16 years and older was based on:

a = Total population of 16 years and older in 1990 of the state of Morelos is 711,775
b = Total population of 16 years and older in 1990 of the state of Morelos without junior high school is 517,488
c = Ratio of educational achievement of the state of Morelos of 16 years and older

$$c = 1 - \frac{b}{a} = 1 - \frac{517488}{711775} = 0.27296$$

State	Poverty Threshold (current pesos)	Life Expectancy Privation Index	Education Privation Index	Economic Privation Index	Average Privation Index	Human Development Index (HDI)
Oaxaca	2,215,584.39	1.00	0.94	1.0000	0.98	0.02
Chiapas	2,364,960.43	0.64	1.00	0.8706	0.84	0.16
Guerrero	2,805,454.17	0.56	0.85	0.4889	0.63	0.37
Puebla	2,937,790.91	0.68	0.65	0.3743	0.57	0.43
Michoacán	2,759,174.72	0.45	0.61	0.5290	0.53	0.47
Hidalgo	3,366,173.22	0.68	0.70	0.0032	0.46	0.54
Tlaxcala	2,929,986.06	0.54	0.42	0.3811	0.45	0.55
Guanajuato	3,237,826.21	0.40	0.62	0.1143	0.38	0.62
Veracruz	3,366,274.25	0.40	0.64	0.0031	0.35	0.65
Querétaro	3,367,844.88	0.48	0.52	0.0017	0.34	0.66
Tabasco	3,367,806.99	0.53	0.47	0.0017	0.34	0.66
Zacatecas	3,141,711.78	0.23	0.49	0.1976	0.31	0.69
Yucatán	3,339,149.93	0.28	0.55	0.0266	0.29	0.71
San Luis Potosí	3,366,363.60	0.28	0.54	0.0030	0.27	0.73
Colima	3,367,839.38	0.39	0.33	0.0017	0.24	0.76
Campeche	3,369,809.99	0.19	0.53	-	0.24	0.76
Morelos	3,367,183.34	0.29	0.38	0.0023	0.23	0.77
Quintana Roo	3,368,448.86	0.23	0.43	0.0012	0.22	0.78
Nayarit	3,365,959.15	0.24	0.41	0.0033	0.22	0.78
Chihuahua	3,368,093.07	0.36	0.28	0.0015	0.21	0.79
Edo. de México	3,367,550.71	0.33	0.31	0.0020	0.21	0.79
Jalisco	3,367,746.00	0.27	0.35	0.0018	0.21	0.79
Aguascalientes	3,366,937.97	0.24	0.29	0.0025	0.18	0.82
Sonora	3,368,567.66	0.32	0.19	0.0011	0.17	0.83
Durango	3,366,722.53	0.23	0.26	0.0027	0.17	0.83
Sinaloa	3,366,951.11	0.18	0.31	0.0025	0.17	0.83
Tamaulipas	3,367,748.48	0.22	0.24	0.0018	0.16	0.84
Coahuila	3,368,580.78	0.22	0.21	0.0011	0.14	0.86
Baja California	3,368,620.91	0.28	0.15	0.0010	0.14	0.86
Baja California Sur	3,368,464.24	0.22	0.17	0.0012	0.13	0.87
Nuevo León	3,369,379.91	-	0.11	0.0004	0.04	0.96
D. F.	3,369,557.80	0.07	-	0.0002	0.03	0.97

Table 4-4 Human Development Index per State

The calculation of the literacy rate for the states, using Morelos as an example is the following:

a = Total population of 15 years and older of the state of Morelos for 1990 was 742,676

b = Total literate population of 15 years and older of the state of Morelos for 1990 was 653431

c = Literacy rate of the population of 15 years and older of the state of Morelos in 1990

$$c = \frac{b}{a} = \frac{653431}{742676} = 0.8798$$

In relation to the economic aspect, the GDP per capita needs to be calculated and that calculation is neccesary to obtain the poverty threshold.

a = Gross Domestic Product of the state of Morelos in 1988 was $4,984,828

b = Total population of the state of Morelos for the year of 1990 was 1,195,059

y = Gross Domestic Product per capita for the state of Morelos

y* = Poverty theshold was $3,365,388. This number was based in the minimum wage

W(y) = welfare for the state of Morelos

$$y = \frac{a}{b} = \frac{4984828 * 10^6}{1195059} = 4171198$$

$$w(y) = 3365388 + 2\sqrt{4171198 - 3365388} = 3367183$$

4.4 Discussion

Table 4.1 presents data necessary to calculate the ratio of educational achievement for the population of persons 16 years and older for the year 1990. We find that the states of Oaxaca, Zacatecas and Chiapas present the lowest index. This means that their population are not on track in relation to education. Oaxaca and Chiapas are states with a high indigenous population and Zacatecas is a state whose inhabitants tend to migrate to the U.S.. The best performeing state in this index is D.F.. One of the reasons is that it posseses the infrastructure that facilitates that the population can be more educated than other states; competition is also a fact. The other high performing states are all northern states; Nuevo Leon, Baja California and Baja California Sur. The type of economic activity performed in these states helps the people there tend to prepare themselves better. These states economic activity is mostly service except the case of Nuevo Leon whose main activity is manufacturing. We find that states such as Edo. de Mexico and D.F. can improve relatively easy because of the population density and infrastucture that they possess. In the case of Veracruz, the index is less than 20%. A possible explanation for this phenomena is its main economic activities, harbor activity, principal entrance of imports, fishing, and agriculture. We can see in general that these areas do not demand a high educational level of the workers.

In Table 4.2 we find the calculation of the literacy rate for the population of 15 years and older. In general the literacy rate is relatively high for each of the Mexican states, with 0.69 being the smallest rate for the state of Chiapas. We can not state that the rates are excellent but in the worst of the cases, we can see that 7 out of 10 persons are literate. On the other hand, it is not clear that the literacy rates are measured using the same standards as in developed countries.

In contrast with the variable educational achievement, the literacy rate is relatively high. In the former the highest rate is 0.42% and in the latter the lowest rate is almost 0.70%. We can see here that the population can read and write but that ability is not used to continue studying high school or to obtain a college degree in the majority of cases. If we also compare these two variables, educational achievement and literacy rate, we can find that two of the states Chiapas and Oaxaca obtain the worst rate, and the states of D.F., Nuevo Leon and Baja California obtain the best rates on each of the variables mentioned. This performance starts to give us a clear idea that the states with high performance obtain that level because they are high achievers in all of the variables calculated.

An important issue here in relation with the literacy rate variable is that in general the rates are high because children learn to read in the early stage of their elementary school education. It is important to mention that the numbers express those that know to read and write but not if they are passive iliterate, which means knowing how to read and write but not knowing the real meaning of what is read or written.

We can also conclude that the discriminant power of the variable literacy rate is less than that of educational achievement. This is the reason why the difference between states of the former variable is so small that the contribution to the HDI is almost the same.

Table 4.3 presents a group of three variables that are reported by INEGI and the fourth varaiable is the GDP per capita calculated using the GDP for the year of 1988 and the population of 1990. aAfter several consultations with economists, they recommended to use 1988 GDP instead of the one for 1993 because of all the changes that the Mexican economy underwent.

Life expectancy is an indication of the health of the population. We can see in the numbers of Table 4.3 that the states that present the lowest life expectancy were: Oaxaca, Puebla, Hidalgo, Chiapas and Guerrero. States such as Oaxaca and Chiapas are mainly composed of a majority of indigenous population, who normally are peasants that work their land to feed their families and themselves. The majority of the land which they use to grow corn and beans are not as fertile as needed to be very productive. The amount of health centers and medical doctors available near to their communities is very low, a situation that complicates the supply of health services. So in general, their feeding habits and the health services available affect the life expectancy a great deal, particularly in Oaxaca and Chiapas. One of the reasons that explains low life expectancy is the inmigration from poor neigbhor states, as well as a large number of isolated small indigenous communities.

The states with the highest life expectancy are: Nuevo Leon and D.F., states with a high degree of education as the other variables related with education express and also we find many public health centers such as public and private hospitals with modern technologies that help their population enjoy good health. The highways and communications infrastructure helps a great deal also in that its population can have access to the health centers, hospitals, a situation which is not the same for the states with poor indices of life expectancy.

The GDP for the year of 1988 shows that the states that generated the highest were D.F., Edo. de Mexico, Jalisco, Nuevo Leon and Veracruz. One more time we can observe how states like D.F. and Nuevo Leon are the ones that are the best whatever the index calculated. D.F. with a vast concentration of economic activity and Edo. de Mexico, the state with the highest concentration of industrial activity achieved GDP's of $83,371,073 and $44,511,815 respectively, which represents 2 to 10 times more than the GDP of many of the other states. To some extent it is logical that most of the industrial and service activity tends to concentrate in the D.F. and Edo. de Mexico because satelite firms that are suppliers and offer services to the large firms will tend to be nearby. Another important issue is that the population concentration in these two states is more than 18 million inhabitants which is almost one fourth of the entire Mexican population. Therefore, firms that are market oriented will tend to establish nearby.

Considering the HDI, then GDP per capita becomes more important than just the GDP because one is interested in the index at the individual level. This is when the population that achieves the GDP becomes very important, like the case of Edo. de Mexico which after being the second highest state in GDP, when considering the GDP per capita it falls back to the eighteeth place. We can also see that a state like Campeche with a GDP of $8,689,697 with a population of 535,185 inhabitants becomes the state with the highest GDP per capita. The states with the highest GDP per capita are Campeche, D.F., Nuevo Leon and Baja California and the states with the lowest GDP per capita are: Oaxaca, Chiapas, Michoacan and Guerrero. Observing the numbers we can see that D.F. is a state that because of the geographical limits the indices and variables calculated are in its benefit which is not so for Edo. de Mexico. In order to increase GDP per capita for certain states like Tlaxcala, Nayarit, Aguascalientes, Colima, Quintana Roo and Baja California Sur, it could easily be done with the establishment of a couple of medium to large firms because this states have less than a million inhabitants. Finally it can be mentioned that the state of Campeche can achieve a high GDP per capita because of the vast amount of oil which is found on its coast and also because it has a very large floating population.

Table 4.4 presents the different indices used to calculate the HDI plus a column that presents the poverty threshold necessary in the Atkinson equation which is used to calculate the welfare produced by income.

In general we find that the state of Oaxaca presents the worst indices except in the education privation index in which Chiapas presents the worst. The HDI is calculated on a relative basis. This means that depending on the set being used to calculate the index, a state will appear better or worse. It is very important to note that the state of Oaxaca presents an HDI of only 0.02. This is dismal compared to the maximum possible index of 1.00. Chiapas is the second to last state with an index of 0.16, which is very low also. Guerrero is the third lowest state with an index of 0.37, and in the fourth lowest place we find Puebla with an index of 0.43. In general we find 12 states with an index less than 0.70 which in academic terms can be considered as a non passing grade. In the set of states with an index less than 0.70, we do not find northern states, states which can be considered small in relation to population except for the case of Veracruz with a population of more than 6 million inhabitants.

The states with highest indices are all northern states except for the case of D.F, which is the economic and political capital of the country. States like Jalisco and Edo. de Mexico do not present higher indices because the migration to some extent affects their performance, although Nuevo Leon, the state with the second highest index, also presents a high immigration rate but it works there to attract very talented people.

Mexico, as a federal country, leaves the decision making process that defines their internal policy in the education, health and economic arena to eacj state obviously taking into account the overall country policy. In general each state seeks the welfare of each of its inhabitants so the calculation of the HDI could serve to the states government to define where they should invest their economic resources and efforts so that the welfare can be improved. To some extent it seems logical to start investing in the education sector in order to positively affect the rest of the areas such as health and the economics, but in order to improve the HDI, employment generation is the most important point in the agenda in order to improve the welfare of the states and of the country as a whole.

Finally, as a country and even as a state, the policy makers should try to start the welfare improvement in the area that best benefits it inhabitants and in the county where most of the people can benefit from it.

5. Optimal Stratification of Mexican States

5.1 Introduction

In statistics and marketing there is an important tool to obtain information about a population, which is survey sampling. The fundamental objective of statistics is making inferences about a population based on the information contained in a sample. If inferences of a population have to be done, it is important to design the way in which the sample is to be taken. Sample design has two important aspects: a) a selection process, which is the rules and operations to be performed in order that some members of the population are included in the sample and b) an estimation process, which is related to computing the sample statistics, which are sample estimates of the population values.

The objectives of a survey are the following:

a) Definition of survey variables

b) Methods of observation (measurement)

c) Analysis methods

d) Utilization of survey results

c) Desired precision of survey results

5.2 Stratification

To have a clear idea of optimal stratification, first we need to lay the basis of what stratified sampling is. Stratified sampling consist of the following steps (Kish, 1965):

a) Divide entire population of sampling units into distinct subpopulations called strata.

b) Select within each stratum a sample.

c) Compute the stratum mean from each stratum. The stratum means are properly weighted to form a combined estimate for the entire population.
d) Compute the stratum variance from each stratum. The stratum variances are properly weighted to form a combined estimate for the entire population.
These four steps constitute the basis of stratification.

The principal reasons for resorting stratification are:
a) Stratification may be used to decrease the variances of the sample estimates. In proportionate sampling, the sample size selected from each stratum is made proportionate to the population size of the stratum. In disproportionate or optimal allocation, sampling rates are used deliberately in the different strata. The variance can be decreased by increasing the sampling fractions in stata having higher variation.
b) Strata may be formed to employ different methods and procedures within them.
c) Strata may be established because the subpopulations within them are also designated as domains of study.

5.3 Optimum Statification in Multivariate Sampling

Optimum statification is concerned with precision and it is aimed at defining a stratification criterion that provides efficient estimators (Jarque, 1981). In general it is assumed that a stratified sample of size n is obtained from a population of size N, and the first parameter to be estimated is the K population means $q_1, \ldots\ldots q_k$ of a set of variables $X_1, X_2, X_3, \ldots\ldots X_k$. If the number of strata = L and N_h = size of the stratum h, then $\overline{\overline{X_{k,h}}}$ = sample mean of variable X_k for stratum h, $\overline{q_k}$ = estimator of q_k using stratified sampling $s_{k,h}^2$ = variance of variable X_k within stratum h and W_k is defined as N_h/N for h = 1, 2, ...,L.

$$\overline{q_k} = \sum_{h=1}^{L} W_h \overline{X_{k,h}} \qquad k = 1,2,\ldots\ldots K$$

and its variance

$$V(\overline{q_k}) = \sum_{h=1}^{L} W_h^{\,2} V(X_{k,h}) \qquad k = 1,2,\ldots\ldots K$$

The problem of univariate optimum stratification can be stated as finding the strata boundaries $x^{(1)}, x^{(2)}, \ldots\ldots x^{(L-1)}$ subject to $x^{(0)} < x^{(1)} < \ldots\ldots < x^{(L-1)} < x^{(L)}$ where $x^{(0)} =$ min $\{x\}$ and $x^{(L)} =$ max $\{x\}$ such that the $V(\bar{q})$ which is the stratification criterion is minimized. The strata boundaries $x^{(1)}, \ldots\ldots, x^{(L-1)}$ must satisfy the following simultaneous equations proposed by Dalenius:

$$x^{(h)} = \frac{1}{2}\left[\frac{\int_{x^{(h-1)}}^{x^{(h)}} xf(x)dx}{\int_{x^{(h-1)}}^{x^{(h)}} f(x)dx} + \frac{\int_{x^{h}}^{x^{(h+1)}} xf(x)dx}{\int_{x^{(h)}}^{x^{(h+1)}} f(x)dx}\right], \quad h = 1,\ldots\ldots, L-1$$

In general, in multivariate optimum stratification the procedures used attempt to define homogeneous groups according to a specific measure (e.g. distance when using cluster analysis technique). The main concern is to minimize the variance of the estimator, and as for the multivariate case because K variables exist a criterion for multivariate optimum stratification is proposed (Jarque, 1981). The criterion proposed is to find S* such that it minimizes the following equation:

$$F(S) = \sum_{k=1}^{K} d_k(S)$$

where

$$d_k(S) = \frac{V_s(\overline{q_k})}{V^*(\overline{q_k})} \quad k = 1,2,\ldots\ldots K$$

This paragraph explains the way in which the numerator and denominator of the latter equation is determined when applying Dalenius equation individually to each of K variables, it provides K stratifications S_1^*, S_2^*,........$S_k^*$and each of the S_k^* gives respectively a lower bound for the variance of $\overline{q_k}$. Lower bounds are defined by $V^*(\overline{q_1})$,..........., $V^*(\overline{q_K})$ and the variance of q_k when using stratification S as $V_S(\overline{q_k})$.

5.4 Results

INEGI (Instituto Nacional de Estadistica Geografia e Informatica) developed and sells a software named Estratifique based on Jarque's 1981 article related to optimum stratification. The software was used in order to compare the results of the Optimum stratification with the HDI.

When using the software estratifique, there are five ways in which the software stratifies the data: a) based on the socioeconomical level, b) based on education, c) based on occupation, d) based on housing and e) based on urbanization.

Two ways of stratification were studied, one based on socioeconomical level and one based on education.

The stratification results based on the socioeconomical level show that the optimum stratification correspond to seven strata when variables taken into account are the following: a) literate population of 15 years and older, b) proportion of the population between 6 and 11 years old that go to school, c) proportion of the population between 12 and 14 years old that go to school, d) proportion of the population between 15 and 19 years old that go to school, e) average school years of the population, f) proportion of the population that earns less than the minimum wage, g) proportion of the population that earns more than five times the minimum wage, h)proportion of houses with sewage services, i) proportion of houses with drinking water, j) proportion of houses with electricity, k) proportion of houses with one room, l) number occupant per room, among 18 variables.

Table 5.1 shows the seven strata and the states that correspond to each of them, the higher the number of the strata the better in socieconomical terms the state.

Stratum	States
7	D.F., N.L.
6	AGS, BC, BCS, COAH, CHIH, JAL, MEX, MOR, SON, TAMP
5	Q.R.
4	NAY, QRO, SIN, TLAX
3	DGO, GTO, MICH, ZAC
2	CAMP, HGO, PUE, SLP, TAB, VER, YUC
1	CHIS, GRO, OAX

Table 5-1States Stratification based on socioeconomical level.

The stratification results based on the education level show that the optimum stratification correspond to nine strata which variables are the following: a) literate population of 15 years and older, b) proportion of the population between 6 and 11 years old that go to school, c) proportion of the population between 12 and 14 years old that go to school, d) proportion of the population between 15 and 19 years old that go to school, e) average school years of the population, f) proportion of the population of 15 years and olders with education greater than elementary school.

Table 5.2 shows the nine strata and the states that correspond to each of them, the higher the number of the strata the better in socieconomical terms the state.

Stratum	States
9	D.F..
8	BC, N.L.
7	BCS, MEX, SON
6	COAH, MOR, TAMP, TLAX
5	AGS, CHIH, DGO, JAL, QRO, Q.R.
4	COL, NAY, SLP, SIN, TAB
3	CAMP, GRO, HGO, OAX, PUE, VER, YUC
2	GTO, MICH, ZAC
1	CHIS

Table 5-2 States Stratification based on educational level.

5.5 Discussion

In general terms we found that the results obtained when calculating the HDI coincide with the way in which the different stratum are conformed. We find that in the socioeconomical classification the states of NL and DF are the states with the best performance and that CHIS, GRO and OAX are the worst. The situation just explained is the same in the extremes as in the calculation of the HDI except for the case of CAMP (Campeche), a state which has one of the greatest GDP per capita. It is important to keep in mind that the classification made in HDI is over just one variable and in the case of the stratum calculated in this Chapter the classification is based on multiple variables, but although the difference, we find many similarities, which supports the results obtained in the calculation of the HDI.

In the case of the stratum classification obtained for the educational level we also find relation especially at the extremes. We can see that the highest achievers in relation to education DF, BC, BCS and NL are also highest achievers when calculating the literacy rate and the educational achievement for the case of the HDI. Coincidence is not 100% because as was mentioned before, to obtain the stratum several variables are taken into account at the same time. In the other extreme we also find the coincidence of CHIS being one of the states with the worst performance, as ZAC, GRO, HGO and OAX.

In general we find that results obtained from either model are almost the same in the extremes, a situation which makes us confident of either model and also it supports the expected value on those states at the extremes, high and low performers.

If strata is calculated in an univariate form, results are even more alike than in the case of multivariate optimum stratification.

6. Data Envelopment Analysis

This chapter presents the DEA analysis, which calculates the efficiency with which each Mexican state achieves its respective level of human development. Two models were studied and the difference between them were the outputs. In the first case the outputs were adult literacy, school years, life expectancy and total income expressed in per capita terms and in the second case the output was the HDI calculated in the previous chapter. We will find that if we compare two states HDI's, the state with the highest HDI between both will not necessarily achieve a higher efficiency than the one with the lowest HDI. The reasons that explain this phenomena are that the DEA technique focuses on determining efficiency, meanwhile calculating HDI is obtaining an index that is relative to the set being evaluated and also the fact that this index which measures a relative level achieved (effectiveness) not how it was achieved (efficiency). The DEA model used is a BCC output oriented for a cross-sectional analysis for the year of 1990. This model was selected because: a) there is no justification to assume that the variables present a constant return to scale and b) the model presents an output orientation due to the fact that the states should strive to maximize the level of human development attained by their population for any amount of resources they have instead of finding ways to decrease the resources used to attain their current level of human development in order to become efficient. In general, one of the responsibilities of any state is to use the resources they have to achieve a high as possible level of human development.

6.1 Variables

To perform a DEA analysis, we must define input variables and output variables. The determination of the inputs and outputs variables are problematic in not-for profit organizations in general (Sinuany-Stern, Mehrez and Barboy, 1994). In this dissertation, the selection of inputs and outputs, as it was explained in Chapter 4, was selected based on the analysis of human development (Desarrollo Humano, 1992) performed across different countries. Three inputs variables were used in the two models that were run, state total investment per capita, total number of physicians per capita and the total number of school teachers and professors per capita of all the educational levels.

Total investment is the amount of money that the state government invests in infrastructure and promotion for the state in the year of 1990. The components of the investment are the construction and maintenance of: buildings, highways, schools, hospitals, water supply lines, irrigation lines, general public works state promotion. The second input variable is total number of physicians per capita that worked for the national health system, Mexican Social Security Institute (IMSS) and State Workers Security and Social Service Institute (ISSSTE), in that respective state and that were registered by the census in 1990. The last input variable, teachers per capita, is the total amount of teachers of the different educational levels registered by the census of 1990 that taught in public and private schools of the state. The three input variables are associated directly and/or indirectly with the output variables, which are used to obtain the HDI for each of the states. Table 6.1 illustrates the data of the input variables for each of the Mexican states.

State (DMU)	Total Investment per capita	Number of Physicians per capita	Number of Teachers per capita
AGS	0.1408	0.0012	0.0121
BC	0.0593	0.0011	0.0118
BCS	0.0138	0.0017	0.0129
CAMP	0.1327	0.0013	0.0124
COAH	0.0476	0.0013	0.0132
COL	0.0704	0.0016	0.0141
CHIS	0.0773	0.0005	0.0098
CHIH	0.0477	0.0010	0.0108
DF	0.2496	0.0028	0.0162
DGO	0.0429	0.0011	0.0140
MEX	0.0612	0.0008	0.0104
GTO	0.0656	0.0008	0.0098
GRO	0.0757	0.0007	0.0126
HGO	0.0483	0.0009	0.0109
JAL	0.0512	0.0010	0.0108
MICH	0.0256	0.0007	0.0109
MOR	0.1200	0.0009	0.0117
NAY	0.0606	0.0010	0.0140
NL	0.0767	0.0012	0.0129
OAX	0.0269	0.0006	0.0101
PUE	0.0358	0.0007	0.0110
QRO	0.0545	0.0010	0.0112
QR	0.0810	0.0012	0.0103
SLP	0.0419	0.0008	0.0132
SIN	0.0728	0.0009	0.0132
SON	0.0778	0.0011	0.0126
TAB	0.3433	0.0012	0.0111
TAM	0.0908	0.0014	0.0125
TLAX	0.1575	0.0010	0.0124
VER	0.1095	0.0009	0.0115
YUC	0.0372	0.0013	0.0119
ZAC	0.0705	0.0007	0.0113

Table 6-1 Input variables (per capita terms)

The other set of variables used in a DEA model are the output variables. Four variables were used for the per capita run: a) percentage of literate population of 15 years and older, b) the average number of years of formal education of the adult population, c) the expected life expectancy for an inhabitant of a specific state, and d) income per capita. The first variable captures the percentage of the population of 15 years and older that know how to read and write. The second variable explains the average number of years that adult population has spent in formal education. It does not take into account short courses that the Educational Secretary does not endorse. The third variable explains the expected number of years that an inhabitant of a specific state would live. The fourth and final output variable is the per capita population income. This variable measures average income that an inhabitant of a specific state earned in the year of 1990. The output variable for the second DEA model is the HDI.

Table 6.2 shows the output variables used for the per capita and the HDI model. In the first column we find each of the 32 states or DMU's, from the second to the fifth column, we find the output variables used to run the per capita DEA model and in the sixth and last column we find the output variable used to run the HDI model. This second model takes that name because the output variable used is precisely the HDI.

6.2 Sources

The main source of information used to obtain the variables were the annual reports of Instituto Nacional de Estadistica Geografia e Informatica (INEGI). The input variable total investment was obtained from Public Finance for states and municipalities publication for 1990 to 1994 from pages 167 and 204 to 206. The second and third input variables number of physicians and number of teachers is information published also by INEGI in its annuals statistics in the states section in pages 543 for the case of the physicians and from page 555 to 563 for the case of the teachers. The source used by INEGI's publication for the case of the number of physicians is the health information inter-institutional group and president Salinas V government report for the

DMU	Literacy rate	Average school years	Average life expectancy	Per capita income	HDI
AGS	0.93	6.16	71.2	7.25	0.82
BC	0.95	7.55	71.4	11.49	0.86
BCS	0.94	6.93	70.8	8.58	0.87
CAMP	0.84	5.93	71.8	5.52	0.76
COAH	0.94	7.05	71.4	7.61	0.86
COL	0.91	6.29	69.4	8.76	0.76
CHIS	0.70	3.91	66.4	3.59	0.16
CHIH	0.94	6.62	69.8	9.05	0.79
DF	0.96	8.71	73.2	9.71	0.97
DGO	0.96	6.08	71.3	5.37	0.83
MEX	0.91	6.76	70.1	7.40	0.79
GTO	0.83	5.12	69.3	15.45	0.62
GRO	0.73	5.02	67.4	4.15	0.37
HGO	0.79	5.33	66.0	4.59	0.54
JAL	0.91	6.55	70.9	7.62	0.79
MICH	0.82	5.33	68.7	5.34	0.47
MOR	0.88	6.70	70.6	7.45	0.77
NAY	0.88	5.79	71.2	6.90	0.78
NL	0.95	7.81	74.1	9.46	0.96
OAX	0.72	4.45	62.1	3.40	0.02
PUE	0.81	5.51	65.9	4.96	0.43
QRO	0.84	5.22	68.3	7.17	0.66
QR	0.87	5.51	71.3	9.08	0.78
SLP	0.85	5.49	70.8	5.02	0.73
SIN	0.90	6.34	71.9	7.88	0.83
SON	0.94	6.94	70.2	9.10	0.83
TAB	0.87	5.39	67.7	5.47	0.66
TAM	0.93	6.82	71.4	7.18	0.84
TLAX	0.89	6.24	67.6	4.92	0.55
VER	0.82	5.47	69.3	5.32	0.65
YUC	0.84	5.96	70.7	5.60	0.71
ZAC	0.90	4.89	71.3	4.01	0.69

Table 6-2 Outputs variables (per capita measurements and HDI).

case of the total number of teachers. All of the variables were divided by the total population related to the variable for the year of 1990 in order to convert them to a per capita basis.

In the case of the output variables for the per capita model, the literacy rate was obtained from table 21.24 of the annuals statistics of INEGI for the year of 1993 and this information source was the XI General Census of Population and Housing of 1990. The variable average school years of the adult population was obtained from Table 21.34 page 665 of the annual statistics of INEGI for the year of 1993. The average life expectancy was obtained from table 21.9 of the annual statistics of INEGI for the year of 1993. Finally, the last variable, total income, was also obtained from INEGI's annual statistics. The source is the table that presents the number of inhabitants whose ranges of income are based on multiples of the minimum wage.

For the HDI model, the output variable is the result of the calculation made in the previous chapter related with Mexican states HDI.

6.3 Results

Table 6.3 presents the ϕ and the efficiency value for the BCC per capita model and the BCC HDI model. If the cells corresponding to ϕ and efficiency, both have a value of 1 the model defines the specific state as efficient; in the case that the cells are different from 1, the model defines the state as inefficient. The efficiency is the reciprocal of the ϕ value. The first column represents the DMU, second and third columns are the ϕ value and the efficiency value respectively for the per capita model and the fourth and fifth columns are the ϕ value and the efficiency value respectively for the HDI model.

We find that 17 out of the 32 states or DMU's are efficient in the BCC per capita model; we also find that 13 out of the 32 states or DMU's are efficient in the BCC HDI model. An observable phenomena is the fact that the efficiency values for the HDI model are always less than or equal to the efficiency value of the per capita model. The DMU's names that have an asterisk (*) are the ones that are efficient in both models.

	PER CAPITA	MODEL	HDI	MODEL
DMU (State)	ϕ VALUE	EFFICIENCY	ϕ VALUE	EFFICIENCY
AGS	1.017	0.983	1.014	0.905
BC	1.000	1.000	1.009	0.991
BCS *	1.000	1.000	1.000	1.000
CAMP	1.025	0.976	1.218	0.821
COAH	1.009	0.991	1.033	0.968
COL	1.047	0.956	1.251	0.799
CHIS *	1.000	1.000	1.000	1.000
CHIH *	1.000	1.000	1.000	1.000
DF *	1.000	1.000	1.000	1.000
DGO *	1.000	1.000	1.000	1.000
MEX *	1.000	1.000	1.000	1.000
GTO *	1.000	1.000	1.000	1.000
GRO	1.048	0.954	1.865	0.536
HGO	1.068	0.936	1.390	0.720
JAL	1.000	1.000	1.010	0.991
MICH *	1.000	1.000	1.000	1.000
MOR	1.012	0.988	1.081	0.925
NAY	1.018	0.982	1.090	0.917
NL *	1.000	1.000	1.000	1.000
OAX *	1.000	1.000	1.000	1.000
PUE	1.000	1.000	1.209	0.827
QRO	1.044	0.958	1.246	0.802
QR *	1.000	1.000	1.000	1.000
SLP *	1.000	1.000	1.000	1.000
SIN	1.000	1.000	1.003	0.997
SON	1.014	0.987	1.105	0.905
TAB	1.054	0.949	1.296	0.788
TAM	1.021	0.979	1.110	0.901
TLAX	1.055	0.948	1.591	0.629
VER	1.038	0.963	1.281	0.781
YUC	1.009	0.991	1.188	0.842
ZAC *	1.000	1.000	1.000	1.000

Table 6-3 DMU's ϕ and efficiency values.

Tables 6.4 and 6.5 present the values of the reference set for each of the inefficient DMU's. The heading of the columns represent the efficient DMU's which are possible candidates to become a reference element for the inefficient DMU; the rows represent each of the inefficient DMU's. The sum of the values of each row should add up to 1, this is taking into account both tables, 6.4 and 6.5. The value of the cell is the proportion or weight in which the specific efficient DMU, represented by the name of the column, is reference for the inefficient DMU. There are several cases in these tables in which the sum of the row exceeds 1 because of rounding. The reference set represents what in geometry is defined as the facet to which the inefficient DMU is projected. Examples of facet were presented in the Chapter that explains the DEA Technique. It is important to mention that every efficient DMU is reference to itself and sometimes to other inefficient DMU's. Although the DMU's PUE and GTO are efficient they do not appear in any column. The explanation for this is the fact that they were not a reference for any inefficient DMU.

	REF.	SET.										
DMU (State)	BC	BCS	CHIS	CHIH	DF	DGO	MEX	GTO	JAL	MICH	NL	QR
AGS	0.015			0.373							0.612	
CAMP											0.808	0.182
COAH	0.190	0.288				0.219				0.006	0.297	
COL	0.265				0.030	0.204					0.501	
GRO			0.134				0.269					
HGO									0.654	0.215	0.008	
MOR							0.551				0.272	
NAY											0.509	
QRO									0.705	0.078	0.177	
SON	0.597					0.337	0.0005				0.053	
TAB							0.609				0.284	0.107
TAM	0.339			0.013							0.648	
TLAX	0.101					0.397	0.282				0.145	
VER							0.004				0.232	0.167
YUC		0.417							0.436	0.043	0.105	

Table 6-4 Reference sets of BCC per capita model.

	REF.	SET.	
DMU (State)	SLP	SIN	ZAC
AGS			
CAMP			
COAH			
COL			
GRO			0.597
HGO			0.124
MOR		0.045	0.133
NAY	0.456		0.035
QRO			0.040
SON			0.013
TAB			
TAM			
TLAX			0.075
VER			0.597
YUC			

Table 6-5 Continuation of reference sets for BCC per capita model.

	REF.	SET							
DMU	BCS	CHIH	DGO	MEX	GTO	MICH	NL	SLP	ZAC
AGS				0.320			0.680		
BC	0.159		0.014	0.446			0.381		
CAMP				0.200			0.800		
COAH	0.283		0.324	0.023			0.370		
COL	0.100						0.900		
GRO									1.000
HGO		0.554		0.257		0.109		0.080	
JAL	0.095	0.408		0.497					
MOR				0.750			0.250		
NAY			0.279	0.430			0.291		
PUE						0.773			0.227
QRO	0.163		0.035	0.695			0.107		
SIN				0.750			0.250		
SON				0.250			0.750		
TAB				0.720			0.280		
TAM				0.160			0.840		
TLAX				0.500			0.500		
VER				0.750			0.250		
YUC	0.524		0.008	0.403			0.064		

Table 6-6 Reference sets of BCC HDI model.

Table 6.6 presents the reference sets for the inefficient DMU's of the BCC HDI model. As in the two previous tables, the columns represent the efficient DMU's and the lines the inefficient DMU's. The state's DF, CHIS, OAX and QR do not appear on the table because they were reference element only for themselves. In the case of the state ZAC, the sum of the row does not add up to 1 due to the rounding process.

The following tables 6.7 and 6.8, present the optimal values for each inefficient DMU after the radial output augmentation and the inputs were decreased as necessary for a DMU to become efficient or to be projected to the efficient frontier.

DMU	Total Investmet per capita	Number of Physicians per capita	Number of Teachers per capita	Literacy rate	Average school years	Average life expectancy	Per capita income
AGS	0.066	0.001	0.012	0.946	7.362	72.446	9.343
CAMP	0.078	0.001	0.012	0.935	7.368	73.562	9.390
COAH	0.048	0.001	0.013	0.949	7.114	72.051	8.679
COL	0.070	0.001	0.013	0.952	7.416	72.629	9.176
GRO	0.069	0.0007	0.011	0.876	5.261	70.319	4.871
HGO	0.048	0.0009	0.011	0.890	6.092	70.503	6.699
MOR	0.067	0.0009	0.011	0.919	6.779	71.428	7.538
NAY	0.061	0.001	0.013	0.903	6.649	72.496	7.246
QRO	0.054	0.001	0.011	0.910	6.611	71.312	7.625
SON	0.055	0.001	0.013	0.953	7.034	71.149	9.227
TAB	0.068	0.0009	0.011	0.917	6.925	71.365	8.172
TAM	0.070	0.001	0.012	0.950	7.707	72.926	10.146
TLAX	0.057	0.001	0.012	0.939	6.583	71.318	7.062
VER	0.074	0.0009	0.011	0.907	5.679	71.944	6.140
YUC	0.037	0.001	0.012	0.923	6.789	71.351	8.119

Table 6-7 Values of inputs and outputs of BCC per capita model if DMU's were efficient.

DMU	Total Investment per capita	Number of Physicians per capita	Number of Professors per capita	HDI
AGS	0.072	0.001	0.012	0.906
BC	0.059	0.001	0.012	0.868
CAMP	0.074	0.001	0.012	0.926
COAH	0.048	0.001	0.013	0.889
COL	0.070	0.001	0.013	0.951
GRO	0.070	0.0007	0.011	0.690
HGO	0.048	0.0009	0.011	0.750
JAL	0.051	0.0009	0.011	0.798
MOR	0.065	0.0009	0.011	0.832
NAY	0.061	0.001	0.012	0.851
PUE	0.036	0.0007	0.011	0.520
QRO	0.054	0.001	0.011	0.823
SIN	0.065	0.0009	0.011	0.832
SON	0.073	0.001	0.012	0.917
TAB	0.066	0.0009	0.011	0.838
TAM	0.074	0.001	0.012	0.933
TLAX	0.069	0.001	0.012	0.875
VER	0.065	0.0009	0.011	0.832
YUC	0.037	0.001	0.012	0.843

Table 6-8 Values of inputs and output of the BCC HDI model if DMU's were efficient.

The following tables are the results of the sensitivity analysis of the calculations given by the DEA output orientation model applied to the per capita data and to the HDI output data. In general these results give decision makers insight of the substitution rates between the different inputs and outputs used in the model.

Table 6.9 gives the input dual variables and the substitution rates between the inputs of the per capita BCC model. Dual variables represent the amount in which the objective function would improve if a unit of the variable is increased. Cells without a number are due to the fact that the division between inputs is undetermined. Another very important issue here is the fact that efficiency range goes from 0 to 1 and that dual variables exceed the 1, this is due to the situation that the original variables were expressed in per capita terms. In order to get the exact number the per capita number should be taken into account.

	INPUT	DUAL	VARIABLES	SUBSTITUTION	RATES	
DMU	TOTAL INVESTMENT	PHYSICIANS	PROFESSORS	INVEST/PHYS	INVEST/PROF	PHYS/PROF
AGS	0.000	0.000	11.664		0.000	0.000
BC	0.000	0.000	11.503		0.000	0.000
BCS	0.681	0.000	0.000			
CAMP	0.000	0.000	14.999		0.000	0.000
COAH	1.534	100.378	0.000	0.015		
COL	0.052	0.000	0.000			
CHIS	0.000	145.582	29.786	0.000	0.000	4.887
CHIH	2.259	0.000	47.413		0.047	0.000
DF	0.000	0.000	0.000			
DGO	0.030	0.000	0.000			
MEX	1.712	147.814	0.000	0.011		
GTO	0.000	0.000	0.000			
GRO	0.000	466.052	0.000	0.000		
HGO	0.797	46.105	9.019	0.017	0.088	5.111
JAL	0.742	42.918	8.396	0.017	0.088	5.111
MICH	0.766	44.293	8.664	0.017	0.088	5.112
MOR	0.000	172.12	0.000	0.000		
NAY	0.499	72.465	0.000	0.006		
NL	0.479	69.629	0.000	0.006		
OAX	4.077	0.000	139.476		0.029	0.000
PUE	1.750	1502.465	0.000	0.001		
QRO	0.770	44.552	8.715	0.017	0.088	5.112
QR	0.000	0.000	21.351		0.000	0.000
SLP	0.906	37.704	0.000	0.024		
SIN	0.000	192.549	0.000	0.000		
SON	0.000	126.14	0.000	0.000		
TAB	0.000	0.000	21.994		0.000	0.000
TAM	0.000	0.000	11.655		0.000	0.000
TLAX	0.000	112.675	2.446	0.000	0.000	46.065
VER	0.000	38.695	17.685	0.000	0.000	2.188
YUC	0.866	32.511	7.946	0.026	0.010	4.091
ZAC	0.280	25.255	15.567	0.011	0.017	1.622

Table 6-9 Inputs dual variables and sustitution rates for the BCC per capita model

The interpretation of the input substitution rates is the following: if the result of the division, which is the respective column heading, is greater than the unity it means that the policy for the state should decrease a unit of the variable of the numerator instead of a unit of the variable of the denominator. This is so because decreasing a unit of the numerator increases the objective function more than the variable of the denominator does. In the case that the result was the contrary, that is, less than the unity the variable in the denominator is the one that should be decreased. If the result is equal to one it means that decreasing a unit of the numerator or decreasing a unit of the denominator gives the same result.

The Table 6.10 presents the dual input variables, the substitution rates between inputs and the dual output variable for the BCC HDI model. We can observe that the input dual variables are greater than the input dual variables of the per capita model. We can also observe that a pattern in the substitution rate emerges for different states; that pattern will be discussed in the next section.

In theTable 6.11 we can find the output dual variables and the substitution rates between them for the BCC per capita model.

The interpretation of the output substitution rates is the following: if the result of the division, which is the respective column heading, is greater than the unity it means that the social expenditure policy for the state should be to increase a unit of the variable of the numerator instead of a unit of the variable of the denominator. This is so because increasing a unit of the numerator increases the objective function more than the variable of the denominator does. In the case that the result were the contrary, that is, less than one the variable in the denominator is the one that should be decreased. If the result es equal to the unity it means that increasing a unit of the numerator or the denominator gives the same result.

	INPUT	DUAL	VARIABLES	SUBSTITUTION	RATES		OUTPUT DUAL VARIABLE
DMU	TOTAL INVESTMENT	PHYSICIANS	PROFESSORS	INVEST/PHYS	INVEST/PROF	PHYS/PROF	HDI
AGS	0.000	0.000	82.927		0.000	0.000	1.220
BC	3.928	284.787	9.153	0.013	0.429	31.114	1.163
BCS	3.882	281.514	9.048	0.013	0.429	31.113	1.149
CAMP	0.000	0.000	89.474		0.000	0.000	1.316
COAH	3.928	284.787	9.153	0.013	0.429	31.114	1.163
COL	1.883	0.000	0.000				1.316
CHIS	0.000	16562.500	0.000	0.000			6.250
CHIH	21.004	0.000	591.154		0.035	0.000	1.266
DF	0.060	0.000	0.000				1.031
DGO	3.845	266.785	0.000	0.014			1.205
MEX	10.288	606.889	43.792	0.016	0.234	13.858	1.266
GTO	1.551	0.000	468.362		0.003	0.000	1.613
GRO	0.000	7162.162	0.000	0.000			2.703
HGO	15.052	887.856	64.067	0.016	0.234	13.858	1.852
JAL	2.739	0.000	92.433		0.029	0.000	1.266
MICH	26.778	1167.048	0.000	0.022			2.128
MOR	0.000	551.948	0.000	0.000			1.299
NAY	3.749	399.610	0.000	0.009			1.282
NL	3.046	324.683	0.000	0.009			1.042
OAX	1005.888	0.000	29759.568		0.033	0.000	50.000
PUE	11.395	6550.215	0.000	0.001			2.326
QRO	5.118	371.086	11.926	0.013	0.429	31.115	1.515
QR	0.000	0.000	410.256		0.000	0.000	1.282
SLP	17.241	751.387	0.000	0.022			1.370
SIN	0.000	512.048	0.000	0.000			1.205
SON	0.000	512.048	0.000	0.000			1.205
TAB	0.000	0.000	103.030		0.000	0.000	1.515
TAM	0.000	0.000	80.952		0.000	0.000	1.190
TLAX	0.000	772.727	0.000	0.000			1.818
VER	0.000	653.846	0.000	0.000			1.538
YUC	4.757	344.953	11.087	0.013	0.429	31.113	1.408
ZAC	0.000	3840.580	0.000	0.000			1.449

Table 6-10 Inputs and output dual variables and substitution rate for the BCC HDI model.

	OUTPUT	DUAL	VAR		SUBST.	RATE				
DMU	LITERATE	AVERAGE SCHOOL YEARS	LONGEVITY	TOTAL INCOME	LITERA./ AVERAGE	LITERA./ LONGEV.	LITERA./ INCOME	AVERAGE / LONGEV	AVERAGE / INCOME	LONGEV / INCOME
AGS	0.778	0.000	0.004	0.000		194.500		0.000		
BC	0.767	0.000	0.004	0.000		191.750		0.000		
BCS	0.000	0.000	0.013	0.009		0.000	0.000	0.000	0.000	1.444
CAMP	0.000	0.000	0.014	0.000		0.000		0.000		
COAH	0.186	0.022	0.009	0.000	8.455	20.666		2.444		
COL	0.974	0.000	0.001	0.002		974.000	487.000	0.000	0.000	0.500
CHIS	0.000	0.000	0.015	0.000		0.000		0.000		
CHIH	0.000	0.015	0.012	0.003	0.000	0.000	0.000	1.250	5.000	4.000
DF	0.000	0.008	0.011	0.012	0.000	0.000	0.000	0.727	0.667	0.916
DGO	0.802	0.000	0.003	0.000		267.333		0.000		
MEX	0.000	0.043	0.010	0.000	0.000	0.000		4.300		
GTO	0.000	0.000	0.012	0.010		0.000	0.000	0.000	0.000	1.200
GRO	0.000	0.033	0.012	0.000	0.000	0.000		2.750		
HGO	0.000	0.000	0.015	0.000		0.000		0.000		
JAL	0.000	0.000	0.014	0.000		0.000		0.000		
MICH	0.000	0.000	0.015	0.000		0.000		0.000		
MOR	0.000	0.015	0.013	0.001	0.000	0.000	0.000	1.153	15.000	13.000
NAY	0.000	0.000	0.014	0.000		0.000		0.000		
NL	0.000	0.000	0.013	0.000		0.000		0.000		
OAX	0.000	0.000	0.016	0.000		0.000		0.000		
PUE	0.000	0.146	0.003	0.000	0.000	0.000		48.666		
QRO	0.000	0.000	0.015	0.000		0.000		0.000		
QR	0.353	0.000	0.009	0.003		39.222	117.666	0.000	0.000	3.000
SLP	0.000	0.000	0.014	0.000		0.000		0.000		
SIN	0.000	0.012	0.012	0.004	0.000	0.000	0.000	1.000	3.000	3.000
SON	0.744	0.004	0.004	0.0004	186.000	186.000	1860.000	1.000	10.000	10.000
TAB	0.360	0.000	0.010	0.000		36.000		0.000		
TAM	0.777	0.000	0.004	0.000		194.245		0.000		
TLAX	0.799	0.003	0.004	0.000	266.333	199.75		0.750		
VER	0.000	0.003	0.014	0.000	0.000	0.000		0.214		
YUC	0.000	0.000	0.014	0.000		0.000		0.000		
ZAC	0.000	0.000	0.014	0.000		0.000		0.000		

Table 6-11 Output dual variables and substitution rates for the BCC per capita model.

Continuing with the sensitivity analysis of the the DEA model, the Tables below will show the substitution rates between output variables and the input variables for the BCC per capita model.

The interpretation of the output to input substitution rates is the following: if the result of the division, which is the respective column heading, is greater than the unity it means that the social expenditure policy for the state should be to increase a unit of the variable of the numerator (output) instead of a unit of the variable of the denominator (input). This is so because increasing a unit of the numerator increases the objective function more than decreasing a unit of the input or variable of the denominator does. In case the result were the contrary, this is, less than the unity the variable in the denominator is the one that should be decreased instead of augmenting a unit the numerator. If the result is equal to the unity it means that increasing a unit of the output or decreasing a unit of the input gives the same result.

	SUBST.	RATES	OUTPUT	TO	INPUT.			
DMU	LITERATE / INVEST	AVERAGE/ INVEST	LONGEVITY/ INVEST	TOTAL INCOME/ INCOME	LITERA./ PHYSICIAN	AVERAGE/ PHYSICIAN	LONGEVITY/ PHYSICIAN	TOTAL INCOME/ PRHYSICIAN
AGS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BC								
BCS								
CAMP	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
COAH	0.287	0.034	0.013	0.000	18.838	2.228	0.911	0.000
COL					0.000	0.000	0.000	0.000
CHIS								
CHIH								
DF								
DGO								
MEX								
GTO								
GRO	0.000	0.000	0.000	0.000	0.000	16.117	5.861	0.000
HGO	0.000	0.000	0.012	0.000	0.000	0.000	0.738	0.000
JAL								
MICH								
MOR	0.000	0.000	0.000	0.000	0.000	2.612	2.264	0.174
NAY	0.000	0.000	0.007	0.000	0.000	0.000	1.032	0.000
NL								
OAX								
PUE								
QRO	0.000	0.000	0.012	0.000	0.000	0.000	0.697	0.000
QR								
SLP								
SIN								
SON	0.000	0.000	0.000	0.000	95.162	0.511	0.511	0.511
TAB	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TAM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TLAX	0.000	0.000	0.000	0.000	94.978	0.356	0.475	0.000
VER	0.000	0.000	0.000	0.000	0.000	0.120	0.562	0.000
YUC	0.000	0.000	0.012	0.000	0.000	0.000	0.459	0.000
ZAC								

Table 6-12 Substitution rates between outputs and inputs BCC per capita model.

	SUBST.RATES	OUTPUT	TO	INPUT.
DMU	LITERATE / PROF.	AVERAGE/ PROF	LONGEVITY/ PROF.	TOTAL INCOME/ PROF
AGS	9.228	0.000	0.047	0.000
BC				
BCS				
CAMP	0.000	0.000	0.215	0.000
COAH	0.000	0.000	0.000	0.000
COL	0.000	0.000	0.000	0.000
CHIS				
CHIH				
DF				
DGO				
MEX				
GTO				
GRO	0.000	0.000	0.000	0.000
HGO	0.000	0.000	0.144	0.000
JAL				
MICH				
MOR	0.000	0.000	0.000	0.000
NAY	0.000	0.000	0.000	0.000
NL				
OAX				
PUE				
QRO	0.000	0.000	0.136	0.000
QR				
SLP				
SIN				
SON	0.000	0.000	0.000	0.000
TAB	8.345	0.000	0.231	0.000
TAM	9.246	0.000	0.047	0.000
TLAX	2.061	0.007	0.010	0.000
VER	0.000	0.055	0.256	0.000
YUC	0.000	0.000	0.112	0.000
ZAC				

Table 6-13 Continuation of substitution rates between outputs and inputs for the BCC per capita model.

The final Table of this section presents the substitution rates between the output HDI and the inputs. In general it can be observed that many more cells have a value and those values are much greater compared with the Tables of the BCC per capita model.

	SUBST.RATES	OUTPUT	TO INPUT
DMU	HDI/ INVESTMENT	HDI/ PHYSICIANS	HDI/ PROF.
AGS	0.000	0.000	11.692
BC	4.609	334.188	10.740
BCS			
CAMP	0.000	0.000	143.416
COAH	4.719	342.137	10.996
COL	3.100	0.000	0.000
CHIS			
CHIH			
DF			
DGO			
MEX			
GTO			
GRO	0.000	36105.139	0.000
HGO	38.748	2885.589	164.926
JAL	3.502	0.000	118.190
MICH			
MOR	0.000	755.055	0.000
NAY	5.238	558.407	0.000
NL			
OAX			
PUE	32.044	18420.082	0.000
QRO	9.661	700.495	22.512
QR			
SLP			
SIN	0.000	618.868	0.000
SON	0.000	681.804	0.000
TAB	0.000	0.000	198.078
TAM	0.000	0.000	106.929
TLAX	0.000	2235.064	0.000
VER	0.000	1288.193	0.000
YUC	7.957	577.004	18.545
ZAC			

Table 6-14 Substitution rates between output and inputs of the BCC HDI model.

6.4 Discussion

The first issue to discuss here is the fact that efficiency values for the HDI model are always less than or equal to the efficiency values of the per capita model and also the fact that being efficient in the HDI, according to results, guarantees that the state is efficient in the per capita model. Supporting those results we find the fact that when the sum of inputs and outputs tends to be big, the model more efficiently defines DMU's than when the model has a sum of inputs and outputs that tends to be smaller. Another important issue here is that some states become efficient in the per capita model for the fact that they are well endowed (inputs) or perform very well (outputs) in just one or two of the total of inputs or outputs that the model uses. On the contrary, due to the fact that the set of outputs are reduced to just one in the HDI model, where the same outputs as the per capita model are used but through a formula that allows as much as 1/3 for each of its components, education, health or income, being outstanding in just one of the elements does not guarantee that the DMU will become efficient.

In relation to the reference sets tables the most important aspect that is worth discussing is the number of appereances of a efficient DMU and the weight it represents for a specific inefficient DMU. In the case of the per capita model results show that the DMU's NL, BC, MEX and ZAC are the ones that appear the most in the different reference sets, NL does not appear in just one of the cases and in five occassions with weights that exceed 0.5. In other words NL is a state that is worthwhile studying in detail because of its achievements and because it is a state that becomes reference for all other inefficient states.

In relation to the reference sets of the HDI model we find that the most important DMU's are MEX and NL because they do not appear as reference elements of inefficient DMU's in 3 and 4 respectively. We also find that the weights are relatively high for each of the cases in which they are reference elements.Efficient DMU's as DF, CHIS, OAX and QR although efficient are special outliers cases, as the DEA literature defines efficiency, due to the fact that they are no reference element to any inefficient DMU. Finally we find that they are the only reference elements of the set, of 6 out of the 18 inefficient DMU's that the HDI model define.

Presentation of the tables related to DMU's projection to the efficient frontier show us the great effort that the governments of inefficient DMU's should make in two directions, a) output augmentation and b) input decrease due to slackness in the usage of their resources. We find that the models give a clear signal to governments in which direction the efforts should be made so that each of the inhabitants of the specific state can be better off due to the increase of the outputs in either of the models. The best way to improve is to know what should be fixed and how much it needs to to be fixed.

In relation to the input substitution rates for the BCC per capita model we find that states as DF and GTO have all of its dual variables equal to zero, so no substitution rates apply to them. Dual variables also show that for states such as AGS, BC, CAMP, QR, TAB and TAM, the only way that they can improve through inputs is affecting the Professors variables; on the other hand, we find that states such as COAH, MEX, PUE and SLP Professors is the only one of the input variables that if we alter them they do not affect the efficiency. This situation shows a clear path to improve efficiency in these states. In one case it expresses what variable should be affected and in the other case what variable, although affected, does not produce any benefit or efficiency increase.

Following the same idea as the paragraph above, we find that states such as MOR, SIN, SON and GRO will only be affected if the input variable physicians is altered; otherwise, no positive changes in efficiency will result. The same idea operates for the states of BCS, COL and DGO in relation to the input variable total investment.

In relation to the substitution rates between inputs in the BCC per capita model we find that HGO, JAL, MICH and QRO follow the same pattern because each substitution rate is exactly the same; that is, in first place decrease physicians rather than decrease professors, in second place decrease professors rather than decrease total investment and at last decrease physicians rather than decrease total investment.

We find that in the case of the dual variables and inputs substitution rates for the BCC HDI model, states as AGS, CAMP, QR, TAB, and TAM the only possible way to increase their efficiency is by affecting the professors variable, CHIS, GRO, MOR, SIN, SON, TLAX, VER and ZAC by affecting the physicians variable and DF the only state that would be affected by total investment variable.

In relation to the substitution rates we find exactly the same pattern in the following states: BC, BCS, COAH, QRO and YUC, that is, in first place decrease a physician instead of a professor, in second place decrease a professor instead of the total investment and in third place decrease a physician instead of the total investment.

If the emphasis of the analysis is placed in the outputs augmentation, the sensitivity analysis for the BCC per capita model shows us that longevity is the only variable that can help the following states to increase their efficiency: CAMP, CHIS, HGO, MICH, NAY, NL, OAX, QRO, SLP YUC and ZAC; all the other states could be affected by at least two of the output variables.

Related to the substitution rates, no clear pattern is visible and due to the fact that there are so few cells with substitution rate the recomendation would be to analyze each state to observe the order in which substitution rates should be done in order to increase efficiency. In general we can state that the substitution rate which gives the highest values is the one related to literacy rate and longevity. The ratio suggests that it is better to make the effort and let inhabitants become literate instead of trying to increase their longevity.

For the case of the BCC HDI model we find that no deep analysis can be made due to the fact that only one output variable exists, a situation that does not allow us to perform substitution rates. But the important conclusion that can be extracted from the results is the excessively high dual variable value that the state of OAX presents, 50, CHIS in second place with 6.25 and GRO with 2.703. The three of them are the worst performers in what to HDI it refers. The signal for the policy makers is that a great effort has to be made in relation to those states as well as PUE and MICH in order to increase their HDI and consequently the welfare of their inhabitants.

In relation to the substitution rates between inputs and outputs for the BCC per capita model only we find that only six DMU's allow some kind of substitution between the outputs and total investment, HGO, NAY, QRO and YUC only with longevity and COAH and COL with three of the the outputs. COAH does allow exchange with all its outputs except with average school years and COL does allow exchange with all its outputs except with total income.

Substitution rates related with the physicians income permits more interaction than the other two inputs, total investment and professors. SON is the only state which has the possibility of substitution between all of its outputs and the physician input. The policy for this state should be first to substitute between literacy and physicians, that is to increase literacy rate and decrease physicians; with the other variables due to the fact that the result of the division is the same decrease physicians rather than increase average school years, longevity or total income. DMU's as AGS and CAMP have a value of zero in all their substitution rates a situation that defines a policy of decreasing physicians instead of augmenting any of the outputs. DMU's as HGO, NAY, QRO and YUC have a positive value in the substitution rate of longevity with physicians where all of them should pursue the policy of decreasing physicians first instead of augmenting longevity first, except for the case of QRO whose suggested policy is the contrary.

The substitution rate with the professors input presents a policy of decreasing professors first instead of increasing total income first; this is valid for all the states that have a zero value in the corresponding cell. The substitution rate with longevity and professors is the one that presents more cells values different than zero but because all of them are less than the unity the policy that should be pursued is decrease professors in first place rather than augmenting any of the outputs. The only states where it is better to increase outputs instead of decreasing the professors input are AGS, TAB, TAM and TLAX. For all the the states that allow substitution between outputs and professors input the policy is to reduce inputs first instead of augmenting outputs.

For the case of the BCC HDI model just three substitution rates can be calculated. The following states do not permit any substitution rate because of their dual variables results: BCS, CHIS, CHIH, DF, DGO, MEX, GTO, MICH, NL, OAX, QR, SLP and ZAC. Another important fact is that all the cells that have a value different from zero are greater than the unity so the conclusion for those DMU's is to increase its HDI before reducing any of the inputs, being the contrary for the case in which the cells value equal zero.

7. Conclusions

In this final Chapter the general conclusions of this dissertation will be presented, covering the three techniques used and the general results that can be extracted. This work is a combination of several techniques, HDI, Optimum Stratification and DEA BCC output orientation model, that can be used by decision makers in order to help them by providing information to take goods course of action. This work can be very helpful for public policy makers at country level or at state level.

7.1 Summary Contributions and Conclusions

We found using the UN technique to calculate Mexican states HDI very interesting. Our results present face value in the sense that the expectation about the states is confirmed, in that high performers were DF and NL, and low performers were OAX, CHIS and GRO. The HDI technique has the limitation that it does not offer an option through which policy makers can define the order in which they should act in order to improve the index. HDI can be conceptualized as a method used to calculate the relative effectiveness with which a state or a country is achieving its development. Important to mention is that the measure is relative and a state can be classified as outstanding but among a set that are low performers; in this sense it could be important to include in a set a virtual state that in relation to international standards is a high performer, which will give us a better view of the reality of the different states performance on an international basis. Taking into consideration that a large literature review was done it is important to note that the calculation of the HDI at a state level has not been done before according to the databases and books consulted, so this is an important conribution in order to know in relative terms how Mexican states are performing in relation to their development. The results help to detect the inefficiencies which become an opportunity for improvement in the areas of education, health and population income. This could help the central government develop their investment project in the areas that are considered for improving the HDI, so a sound national program related to education can be developed or for health or income, (e.g. high priority educational programs states or high priority income improvement states).

We can also conclude from the results that states such as Campeche, although it is an outstanding performer in the GDP per capita, due to its natural resources, in the calculation of its HDI it is not as good as in the GDP case. This situation should be taken into account by other states in the sense that being good in one element is not enough or a guarantee to have a well developed state; another important conclusion is that because you are well endowed with a natural resources does not necessarily mean you will be a well developed state. The analysis showed us both extremes, well endowed states that are not high performers and also states that are not well endowed with natural resources that are high performers in their HDI, e.g. NL. The HDI technique, as was mentioned before, does not permit policy and decision makers decide the order in which variables should be improved. Taking as an example what was observed in CAMP, then country or state decision makers will follow different patterns in trying to foster development, e.g. simultaneously improving education, health and income to have a sound state development, or improving education due to the fact that it has direct impact on the other variables. HDI technique should be improved in order to offer the best pattern that a state should follow to improve its development situation.

When calculating and adjusting the HDI we also calculated the index excluding the oil effect in the GDP for states as CAMP, CHIS, NL, SLP, TAB, TAMP and VER the results were almost the same except for the cases of DF that increased its HDI by 0.01 and CAMP that decreased its HDI by 0.03 units.

The application of the optimum stratification gave a classification which was very similar to the order which the states HDI presented, especially in the extremes. One of the reasons is that the source of information is the same and the outliers or extreme points are easily captured by either method. Results of applying optimum stratification are supportive for results obtained by HDI and vice versa, even more in the case of univariate stratification calculated with the Estratifique software developed by INEGI.

One important conclusion is that the HDI methodology and the DEA technique used in combination is a very good combination in order to check the efficiency with which the development level is obtained. In the case of states as Chiapas we conclude that although it is a low HDI performer it is relatively efficient in the way it uses its inputs. States such as DF, Baja California Sur (BCS) and Guanajuato (GTO) are robustly relative efficient, which means that even if the data was changed from per capita outputs to HDI output they were always efficient. Those are examples to study in detail in order to learn of the sound relationship they have between their inputs and their outputs.

To become a high performer in either HDI or DEA the state should have a balance in relation to health, education and income between the rural and non rural areas. States such as Puebla, although its capital city is relatively well developed, has many rural communities that are very far away of being considered as developed communities.

We also found that education plays a very important role for states such as DF and NL, situation that also helps to attract talented people who earn very good wages, which is not the case for states such as Oaxaca, Guerrero, Hidalgo and Chiapas, which naturally well endowed but not taking the advantages of its natural resources.

When DEA was run with the per capita inputs and outputs it was found that the per capita numbers related to physicians and teachers where more or less similar, a situation that made us think that government resources are allocated taking into account the population as a parameter.

An important issue is the fact that we obtained efficient states of the northern, central and southern part of the country from which we can conclude that geography is not the main variable for efficiency or inefficiency.

We also found that DEA results in this dissertation support Bannister and Stolp's 1995 paper conclusions that although increasing cost of congestion regions that are more heavily industrialized exhibit higher levels of efficiency. Among those states were DF, NL, JAL and Mexico, which are a subset of the efficient states obtained for the HDI.

7.2 Limitations

As any research which is done, this dissertation has several limitations that should be taken into account for the interpretation of the results. The first limitation is one related to the data. First of all, the GDP used in the study for calculating the poverty threshold is for the year of 1988 instead of 1990 for which it was not calculated. The data used to calculate the HDI are not exactly the same as the one used by the UNDP to calculate the index at the world wide level, but they are aproximately the same. The validity of the data is as valid as the data collected by INEGI in their census for the year of 1990. All the conclusions drawn for this research are valid for the year of 1990 which the data used represents.

7.3 Directions for Future Research

One important step ahead could be to study the HDI within each state, where each delegation for the case of the DF, or each municipality or county becomes the DMU so that one could know which counties have priority to improve their development level in a DEA efficient manner. It would also be interesting to replicate the study for the year 2000 given that the data used will be obtained again by the census of that year.

Another importan step for future research could be changing the input variables and or running DEA models with absolute variables in order to observe the efficiency results so as to compare them with the results of the per capita models and capture the relationship between both.

Another direction could be to focus the research to an input orientation in relation to DEA. This would mean to obtain valuable information on which of the inputs are not used to a full capacity, not to decrease states inputs but just to observe states under usage of their inputs.

A possible extension would be exploiting returns to scale possibilities, so that it could be known which resources give more than proportional increases.

Finally it could be benefical conducting a study to observe the impact of the neighbors states on the efficiency of each specific state.

8. Bibliography

Andersen, P. & Petersen, N.C.. "A procedure for Ranking Efficient Units in Data Envelopment Analysis." Management Science, Vol. 39, No. 10 October 1993: pp. 1261-1264.

Bannister, G.J.& Stolp, C.. "Regional concentration and Efficiency in Mexican Manufacturing." European Journal of Operational Research, Vol. 80, 1995: pp.672-690.

Bartol, K. M.; Martin, D. C.; Management. McGraw-Hill 1994.

Charnes, A.; Cooper, W.W.; Lewin, A. & Seiford, L. M.; Data Envelopment Analysis, Theory, Methodology and Applications, Kluwer Academic Publishers, 1994.

Charnes, A,. Cooper, W.W. & Li, S.. "Using DEA to Evaluate the Efficiency of Economic Performance by Chinese Cities." Socio-Economic Planning Science, Vol 23, No. 6 , 1989: pp.325-344.

Charnes, A.; Cooper, W.W. & Rhodes, E.. "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through." Management Science, Vol. 27, No. 6 June 1981: pp. 668-697.

Desarrollo Humano; Tercer Mundo Editores; Santafé de Bogotá, Colombia; 1992.

Evans, J. R.; Anderson, D. R.; Sweeney, D. J.; Williams, T. A.; Applied Production & Operations Management, Second Edition. West Publishing Company, 1987.

Finkler, M. D., & Wirtschafter, D. D.. "Cost-Effectiveness and Data Envelopment Analysis." Health Care Management Review 18, Summer 1993 pp.81-88.

Fogarty, D. W.; Hoffmann, T. R.; Stonebraker, P. W., Production and Operations Management. South-Western Publising Company, 1989.

Golany, B., Phillips, F.Y. & Rousseau, J. J. "Models for Improved Effectiveness Based on DEA Efficiency Results." IIE Transactions, Vol. 25, No. 6, 1993: pp. 2-10.

Informe Sobre Desarrollo Humano. Por Programa de las Naciones Unidas para el Desarrollo. CIDEAL, Madrid, España, 1993.

Instituto Nacional de Estadística, Geografia e Informática, (INEGI), El Ingreso y el Gasto Publico, Referentes a 1990, Mexico D. F.,1993.

Instituto Nacional de Estadística, Geografia e Informática, (INEGI), Anuario Estadístico de los Estados Unidos Mexicanos, Referentes a 1990, Mexico D. F., 1993.

Instituto Nacional de Estadística, Geografia e Informática, (INEGI), Finanzas Publicas Estatales y Municipales de Mexico 1990 1994, Mexico D.F., 1995 pag. 204-206.

Jarque, C.M.;"A Solution to the Problem of Optimum Stratification in Multivariate Sampling." Journal of the Royal Statistical Society. Serie C. Vol. 30 No. 2, 1981: pp163-169.

Kish, L.; Survey Sampling, John Wiley, 1965.

Ley Federal del Trabajo y Leyes Reglamentos E Instructivos Complementarios. Berbera Editores, México 1994.

Retzlaff-Roberts, D. L.. “Relating Discriminant Analysis and Data Envelopment Analysis to One Another.” Computers and Operations Research, Vol. 23, No. 4, 1996: pp. 311-322.

Samuelson, P.A., Nordhaus, W. D.. Economia. Mc Graw Hill 1993.

Schefczyk, M.. “Operational Performance of Airlines: An Extension of Traditional Measurement Paradigm.” Strategic Management Journal, Vol 14, 1993: pp.301-317.

Seiford, L. M.. “A Bibliography of Data Envelopment Analysis (1978 - 1990), version 5.0, 1990.

Sinuany-Stern, Z.; Mehrez, A. and Barboy, A.. "Academic Departments Efficiency Via DEA." Computers and Operations Research, Vol.21, No. 5, 1994: pp. 543-556.

Stevens, J.. Applied Multivariate Statistics for the Social Sciences. Lawrence Erlbaum Associates, 1986.

Sueyoshi, T. "Measuring Technical, Allocative an Overall Efficiencies Using a DEA Algorithm." Journal of Operational Research Society, Vol. 43, No. 2, 1992: pp. 141-155.

Sumanth, D. J.; Productivity Engineering and Management, McGraw-Hill, 1985.

Yue, P. "Data Envelopment Analysis and Commercial Bank Performance: A Primer with Applications to Missouri Banks." Federal Reserve Bank of St. Louis Review, Vol. 74 January-February 1992: pp.31-45.

9. Vita

Didimo Antonio Dewar Valdelamar was born in Colon, Panama, on April 30, 1959, son of Horacio Constantino Dewar and Celia Emilia Valdelamar Achurra. He obtained a BS.degree in Industrial and Systems Engineering at Instituto Tecnologico y de Estudios Superiores de Monterrey (ITESM), Campus Monterrey in December of 1982 and he obtained a Master of Business Administration degree at ITESM, Campus Morelos in December of 1987.

He has worked in several manufacturing firms in Panama and El Salvador. In 1985 he became a faculty member of ITESM, Campus Morelos where he has collaborated in several positions, as Director of the Industrial and Systems Engineering Department, in the Quality and Productivity Center of the Graduate and Research Division, and Dean of the Business and Social Sciences Division, the position he currentlly hold.

In 1990 he started his doctoral studies in the Business Program that was offered by ITESM, Campus Ciudad de Mexico in collaboration with the University of Texas at Austin.

Address:

Privada de Colorines # 102
Santa María Ahuacatitlán
Cuernavaca, Morelos 62100
México

This report was typed by the author.