

Bacares, Carlos Augusto Olarte

Conference Paper

Do public transport improvements increase employment and income in a city?

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Bacares, Carlos Augusto Olarte (2013) : Do public transport improvements increase employment and income in a city?, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124085>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Do public transport improvements increase employment and income in a city?

Carlos Augusto Olarte Bacares
Centre d'Economie de la Sorbonne (CES)
University of Panthéon-Sorbonne, Paris 1

March 1, 2013

Abstract

Previous researches have proved the existence of a causal relationship between the concentration of jobs in a city and the income of inhabitants. Other researchers have studied the close and even nearly causal relationship between those variables and the degree of accessibility or of infrastructures like highways on different zones of a city. Nevertheless, no one research has taken into account the degree to which each area of a city benefits from the latest improvements of public transports. The aim of this research is to analyze the relationship between the size of labor market, the income and the employment concentration on different zones of a city with respect to the degree of improvements of public transports (Transmilenio) in a city of a developing country like Bogota. The degree of enhancements of public transports in a zone is suspected to be endogenous to those three variables. Through the use of OLS estimations and then 2SLS estimations for each dependent variable (3), the validation of endogeneity gives enough tools to infer a positive or a negative causality of improvements of public transports. The size of companies defined by the number of jobs they offer plays a role of instrumental variable to explain the degree of improvement of public transports on each zone. It suggests three different scenarios that depend on the size of companies established on each zone. In essence, the number of jobs, the size of the labor market and the income are largely defined by the level of improvements of urban public transports on each zone of the city and the causality relationship changes depending on the size of companies established on each zone. Results of this empirical research give some robust evidences. In the case of Bogotá, public transports improvements have a causality relationship on the income of inhabitants on each zone and on the number of jobs and it change with respect to the size of enterprises. In contrast, the size of labor market, defined as the number of reachable jobs on a specific time, is not a cause of the degree of the presence of public transports' enhancements.

JEL Codes: R 12, R23, R40, J68, J44

Keywords: Causality; Improvements of public transports; endogeneity; Effective size of labor market; Size of enterprises.

1 INTRODUCTION

1.1 A brief theoretical contextualization

Connection between commuting time and localization of jobs in urban areas has been widely studied by several researchers (Cervero 2000; Hansen 1978; Prud'Homme 1999). The number of jobs that can be reached in a specific interval of time, which can be directly related to the concept of job accessibility (Kay Axhausen 2008), represents a relevant topic on urban studies and transportation researches. Time of urban travels, distance from homes to jobs and urban structure seem to hoard the attention of the majority of researchers on urban and urban transports researches.

Inhabitants choose to live in a specific zone of a city taking into account numerous elements of houses and neighborhoods according to the theory of hedonic price (Rosen 1974). The area of the house in square meters, the comfort of the house, the materials which with houses are built, seems to be very important to buyers. However, in the last decades, due to the increasing congestion on cities, inhabitants have been giving an increasing importance to the amenities of the neighborhood and the zone where their houses are placed (Glaeser 1999; Brueckner 2001; Putman 2000). The better and more amenities the neighborhood offers, the expensiveness the house will be. In other words, the actual configuration of cities and thus, the choice of inhabitants about the zone where they decide to live, do not depend just on their desires but more precisely, it depends on the facilities of the zone and hence on the price of the amenities the area offers.

Therefore, it is not strange to read that disparities of income in a city represent one the principal cause of the socio economic gap between different zones of a city. (Glaeser and Thyse, 2004; Cao, Morkarian, Handy 2007). Because of their high purchasing power, people and companies with a high level of income used to have more possibilities to choose the zone of the city they want to dwell; they have more flexibility to move and change their place where they live (Brueckner 2002; Glaeser, Rosenthal, Strange 2010; Anas, Arnott, Small, 1998).

On the 70's or 80's, studies about American cities shown that, because of the success of the private car, people used to prefer to inhabit far from all kind of center of the city (commercial center, job center, government center, etc) and go to suburbs near people with the same level of income and the same level of education. They used to prefer that kind of suburbs with less noise of cars, less congestion, better security and that offer big parks and more spaces of recreation than zones closer to the center of the city (Giuliano, Small 1991, Fujita 99). In addition, as I denote before, those kinds of suburbs have a most marked homogeneity of social classes. Due to high prices of houses, people with small level of income will not be able to live there. This phenomenon was generally perceived both in developed country cities and in cities of developing countries. It can be perceived as a phenomenon of cities in general.

This research is focus on the city of Bogota, the capital of Colombia. As some studies have shown for other cities of developed countries, the city center of Bogota had a decline from the urban and planning perspective. A spatial mismatch began to settle down in the city (Bureckner, Zenou 2003).

1.2 Some characteristics of Bogotá

Before 1960's, provision of public transport was led by public administration. Buses, taxis and lines of tramway composed the urban transport system of the city with the tramway as the "spinal column". The city center was a desirable place to live.

Nevertheless, because of a huge public debt of Bogotá and due to different social factors¹, tramway took ten years to close and public administration decides to privatize public transport system. Since the 60's and due to the destruction of a big part of buildings during several civil manifestations, the center of Bogotá stopped to be a desirable place to live. The management of public transport was given to private entrepreneurs which were just focused on the development of more lines of buses and not on the urbanization or the planning of the city. Public administration was supposed to have a regulator role but, because of the big lobby of private entrepreneurs, that role was not assured by the administration.

Today, public transport system of the city continues to be managed by private entrepreneurs assembled in about 60 private enterprises assuring the provision of the service in the city². Public administration does not have participation on utilities. Since the late 60s, the administration has been taking its role as regulator and determines the different lines of buses needed in some zones of the city and after a call for tender, it assigns those lines to "winner" enterprises. This mode of management contributed to the expansion of public transport system and thereby the big spread of the city on last 40 years.

During those decades downtown became older, not safe, noisy during the day and abandoned and dark during the night. Because of these characteristics, during this period, prices of houses at downtown decrease significantly; people with low income opted to live on the center. However, it does not mean that they live only in the center of the city. They also live in suburbs, far from the center of the city, but with characteristics widely different than those of zones where "rich" people use to inhabit.

That type of behavior encouraged the development of inner cities (Small 99; Fujita 89; Duranton, Anas 2001) and also an increasing expansion of urban areas.

¹Public debt and "El Bogotazo": civil insurrection after the homicide of a very important politician (Jorge Eliecer Gaitan) a future president of Colombia at 1948.

²From Secretaria de Movilidad del Distrito

Generally, suburbs where people of low level of income live do not have a good coverage of public services. The water and sewer services are not provided in the same proportions. Clean water, electricity and even paved roads and therefore, urban transports are not offered in the same proportion.

Population continues to grow rapidly in the urban area and hence, congestion is growing faster than before. Nowadays, people take more time that they are ready to spend from their houses to their jobs. Cities are still to expand its boundaries and “exclusives” suburbs are now too far from the city and what is worst, from work. As a consequence of the sprawl and the big lack of planning policies for decades, Bogotá suffered big problems with mobility and transportation system. Densification of downtown was one of the principal goals of the elected mayor on 2011³. The continuous reduction of accessibility, the sever congestion in some main roads of the city in addition to the large travel time and poor road network conditions, became the most relevant evils of the city. These consequences have been called by some researchers as the “spatial mismatch” (Kain,1968, 1993; Ihlandfeldt and Sjoquist, 1991).

At the end of 1990’s, being aware this big need of the city, the major of Bogotá decided to improved and innovate infrastructure of public by the creation a transport system called “Transmilenio” which has been the most efficient BRT system ever seen in Bogotá (Chaparro, Irma 2000; Hidalgo, Sandoval 2001).

A reorganization of transport network and a restructuration of planning policies are now at the core of the public debate. Today, public administration participates as manager and as inspector; a change of social and urban structure took place in Bogotá on last fifteen years. Most of neighborhoods (60%) are now connected directly by Transmilenio (TM). Most of zones (excluding the suburbs) will be inter-connected substantially with the downtown by TM⁴.

Consequently a shift about the decision to live in suburbs took place. From the last fifteen years, people have been shown an increasing willingness to inhabit the center of the city because of the proximity to their jobs. In Bogotá, more than 30% of jobs are located in the downtown. This fact had as consequence a “redensification” phenomenon of the city center by people of highest income as well as a lift up of prices in the center. Actually, people prefer to live near to their jobs and not at an hour and a half from them⁵. Thus, the connection of income, job accessibility, employment and public transport improvements seems to be very close.

Some earlier studies (Immerluck 98; Gao, 2006; Gao, Mokhtarian, Johnston 2008), have demonstrated the dependence and even the causal relation among accessibility to jobs, employment, income and auto-ownership. In this study,

³Gustavo Petro was elected for the period 2012-2016

⁴Source: Secretaria de movilidad del Distrito

⁵Time that they normally take if they live in suburbs. Encuesta de movilidad del distrito 2005. Secretaria de movilidad del distrito

I propose to demonstrate differences of these relationships among zones benefited with the improvements of transport network and those that do not had enhancements.

Some researches about hedonic prices and its relationship with public transports reveal that the distance between houses to public transports networks can have positives or negatives effects on the hedonic prices of houses⁶. Regarding the city of Bogotá, Mendieta and Perdomo proved that “the average elasticity proximity of TM, price of the land are -0.36, -0.55 and -1.13 for the first, second and third stages of TM in the same order”⁷. In parallel, Gutierrez (2011) suggests that amenities of zones, the accessibility, economies of scales and other characteristics of zones contribute to explain the configuration of employment centers.

Taking into consideration those studies and following some theoretical and empirical studies (Gao, 2006; Gao, Mokhtarian, Johnston 2008) this work expects to find that the presence of improvements of public transports on each zone of Bogota has a significant impact on the level of income, the number of jobs and the level of what Prud’Homme and Lee (1999) called “The Effective Size of Labor Market” (*ESLM*) one each zone. Differences of income, employment and *ESLM* between zones with improvements of transport networks and zones without the presence of Transmilenio, is expected. This research tends to prove that these three variables are not only deep related to the improvement of public transport but that there exists an important endogeneity between the level of enhancements of public transports on these three variables. If endogeneity among improvements and other variables is verified, it will let us infer the existence of a causal relation from improvements to income, employment and *ESLM*.

This paper is structured in three main sections, as follows: we first explain the research methodology of the conceptual framework of the subject. We define our instrumental variable (*IV*) and thus, the instruments we use on our models three models. We then introduce the core phase of our analysis that focuses on the demonstration of endogeneity. It may show that there is a causal relation between these factors. In fact, from a theoretical point of view, an endogenous variable is a factor which is almost determined by the importance of other variables in a specific model. Generally, factors can be affected directly by other factors presented in the model but also by factors that could not be included in the system. It means that we can find some degrees of endogeneity (partial or total endogeneity).

As we already said, our purpose is to find and reaffirm the theoretical endogeneity between improvements of public transport system and job accessibility,

⁶In the case of Bogota the distance is measure with respect to the nearest BRT station, Transmilenio.

⁷Mendieta, J. C. y Perdomo, J. A. (2007). “Especificación y estimación de un modelo de precios hedónico espacial para evaluar el impacto de Transmilenio sobre el valor de la propiedad en Bogotá”, Documento CEDE, Centro de Estudios sobre Desarrollo Económico, 22, ISSN 1657-5334

income and employment⁸. Finally, in order to better understand the model' results, a discussion take place in a fourth chapter. By this way we will be able to conclude if a causal relationship between these variables is plausible.

2 CONCEPTUAL FRAMEWORK AND RESEARCH METHODOLOGY

Earlier studies (Thompson 1997; Sanchez 1999; Shen, Sanchez 2005) have shown the relationship between employment, income, job accessibility and auto ownership. This research assumes this relationship and it aims to demonstrate the importance of the enhancement of public transports provided by the construction of TM, on the level of income, on the number of jobs and on the degree of accessibility to jobs defined by the *ESLM*.

According to Kawabata (2003), interdependence between these variables can be expected. In effect, the improvement of urban transports influences directly the index of job accessibility. In the other hand, an upgrading of job accessibility will boost the average income because of the gain of time and the reduction of commuting costs which will have also a direct effect on the number of jobs. Indeed, the reduction of commuting costs and on time of travel will brings inhabitants closer from their employments (Prud'homme, Lee , 1999)⁹. On this study, it is assume that this proximity and the reduction of commuting time is due to the presence of more stations of Transmilenio.

2.1 How to mediate with endogeneity

To demonstrate if there is a causal relation, a deep study about the endogeneity of those variables must be done. First, an OLS model for each variable with respect to the others control variables will be done in order to know the degree of relationship between those variables when they are dependent or independent factors. First step will be an examination of the correlations and the variance inflation factors for each variable for each model in order to determine if the model could have some problems of multicollinearity.

After doing those analyzes, the research will proceed with the measure of the possible causal effects. Actually, when causality is evoked, the three possible reasons proposed by Kenny (1979) have to be considered:

- Factor x must precede temporally the dependent variable (not always)

⁸In the model employment will be taken as the total of jobs on each zone of the city.

⁹See the analysis about the relation between the Speed, the Sprawl and the Spread.

- Factor x is correlated with dependent variable
- The relation between the independent and the dependent variable must not explained by other causes.

Condition 1 can be necessary but is not a sufficient condition to establish causality. Even so, due to our type of data, we will not demonstrate this condition because we are not doing a time series analysis. This analysis is statistic, hence, factor x will not precede temporally the dependent variable. Condition 2 needs, from a statistical point of view, a relationship between explanatory variable and dependent variable (which is supposed it was shown before). Third condition supposes independent variable and dependent variable have a relation that depends to other reasons. Then, the independent variable is endogenous and the coefficient of this variable does not reflect a simple correlation. In addition, dependent variable could also explain independent variable (reverse causality). To summarize, endogeneity of variables means that there is a presence of functions explanatory variables in the system.

According to Antonakis et al (2010), independent variables might be endogenous because of many reasons. Endogeneity may be caused by the omission of some variables that explain an independent variable. Finally, the fact that independent and dependent variable may be collected from the same rating source, could also be a reason of endogeneity among other reason we will not discuss in this issue.

From a mathematical point of view, one or some explanatory variables of a model may be endogenous because they can be correlated with error term of the system (condition 3). If it is the case, the least square estimator found is not consistent. To have a consistent estimator, researchers have to find valid instruments of the variable suspected to be endogenous. Subsequently, using the appropriate variables, a two-stage least squares (2SLS) model will gives a better and consistent estimator that will reflect to “real” influence of the endogenous variable on the explained variable.

After doing the 2SLS analysis, the analysis will follow with a test and a correction of the potential endogeneity. Then, it is supposed that results will expose enough tools to conclude if there is endogeneity between variables and thus, if a causal relation is verified or not.

2.2 Models and variables structure

As it was noticed before, it is assume that there is a relationship between the enhancement of public transports provided by the construction of TM, the level of income, the number of jobs and the degree of accessibility to jobs. Note that there are three dependent variables, (three different models to test), and one variable suspected to be endogenous.

2.2.1 Endogenous variable and choice of instrumental variables (*IV*)

Improvement of public transports ($Improvement_i$) will be the “potential” endogenous variable and it will be denoted by the number of stations of Transmilenio on each zone; ($Improvement_i$) = 0 means that there is no presence of Transmilenio on the zone.

The choice of the instruments was made after a simple GIS analysis about the configuration of dependent variables and the instruments. Annexes 6 and 7 show that even if the working force has a “suspected” degree of relationship with employment and income (figure 1 and annex 4), their scattering on the city seems to be randomize and thus, there is not a big correlation between those variables. In effect, it can be seen that there are zones where the employment is condensed but not for working force and vice versa. The same case is perceived concerning the size of enterprises and independent variables. The concentration of enterprises and their size seems to be random and there is not a direct relation with income, the *ESLM* and the number of jobs. Analogically, these remarks could be join with the conclusions of Wenglenski (2006) with respect to the concentration of jobs and the degree of accessibility equality for people of low income in the way that there is not a clear relationship about the concentration of these elements in the city.

2.2.2 The model: three studied scenarios

This research tries to adapt the model presented on some precedent studies (Cao, Mokhatarian, Handy 2007 Cervero 2003) with available data collected for the city of Bogotá. It is assumed that the presence of a station of Transmilenio on each zone depends on the working force of each zone i , WF_i and on the number of enterprises on each zone $Enter_{si}$ with respect to their size s . Variables like employment, income, area and other ones that could be taken in consideration are supposed to be exogenous in order to avoid multicollinearity problems on main models. Remember that IV tries to capture effects of exogenous variables that are not taken in consideration on the “main” models, so on dependent variables. But what were the reasons to choose instrumental variables?

Regarding the number of enterprises, $Enter_{si}$, because of the data available, I decided to classify them on three different groups according to the number of employees and also taking in consideration some details of the methodology from the Colombian government¹⁰. Enterprises with less than ten employees are named “microenterprises” (*MiE*). Enterprises with more than ten employees and less than 50 employees are called small and middle-size enterprises (*SME*). Finally, enterprises with more than 50 employees are considered as large enterprises (*LE*). The reason to differentiate companies from each other’s regarding

¹⁰DANE, Departamento Administrativo Nacional de Estadística in spanish

its size is to take in consideration the number of enterprises on each zone but also the potential jobs that those companies offer. For each model, three different results will be propose with respect to the number of companies which are classified on three groups in accord to the number of employees they have.

Income, employment and accessibility to jobs are the dependent variables to regress.

In this study, accessibility to jobs will be denoted by the Effective Size of Labor Market *ESLM*, which is the number of reachable jobs for people commuting from one zone to another one in a specific time. The *ESLM* was already computed on a previous research¹¹ following the methodology of Prud'homme, Lee (1999). Results are given on three different intervals of time (40 minutes, 50 minutes and 60 minutes)¹² thus, I test three different levels of *ESLM* hence three different models, one per interval of time. A deeper description of this variable is formulated on section 3.

The variable suspected to be endogenous is defined as follow:

$$impro = \pi_0 + \pi_1 Dens_i + \pi_2 WF_i + \pi_3 Enter_{si} + u \quad (1)$$

Models and relationships I propose are:

Model 1 :

$$ESLM_{ti} = \beta_0 Cons + \beta_1 impro_i + \beta_2 Y_i + \beta_3 Pop_i + \beta_4 Area_i + \beta_5 Dist_i + \beta_6 Car_i + \beta_7 X_i + e \quad (2)$$

Where:

Cons, is the constant variable of the model. *ESLM_{ti}*, designates the number of reachable jobs for people commuting from the zone i in a specific time t. It is the “proxy” of job accessibility. *Impro_i*, denotes the number of stations of Transmilenio on each zone of the city. The more stations the zone has, the better will be the improvement of public transports in these zones; no presence

¹¹Olarte, C. (2012) “Heterogeneity of social classes and job accessibility: implications of transport policies in Bogota” Working paper, Centre d’Economie de la Sorbonne (CES).

¹²I suppose that when people take less than 30 minutes to reach their jobs, they prefer do it by walk and not by public transports. In fact, as it is shown on annex 1, the average time of trips done on Transmilenio is 40 minutes and the time of trips done by walk or bicycles is 30 minutes.

of stations in the zone means that there is no improvement or not presence of TM on i . Variable Y_i designates the average level of income for inhabitants living in zone i . $Area_i$, represent the size in hectares of the zone i . $Dist_i$, denotes the average distance that an inhabitant of zone i takes to reach their jobs on t minutes. Car_i , indicates de number of car owners on zone i . X_i , reflects a vector of control variables to complete and to fill in the missing information of the model.

What is suggested with this model is that the size of labor market for people living in zone i depends largely on the degree of improvements of public transports on the zone, the mean income of inhabitants on each zone, the population and the area of each zone. Other variable could also explain the *ESLM* like the mean distance from inhabitants houses to their jobs, the number of car owners on each zone and other control variables represented by vector X_i .

I decided not to take into consideration the variable *emplo* which denotes the number of jobs on each zone because these results (*ESLM*) were made directly with respect to the number of jobs on each zone so, a big problem of collinearity was avoided.

As previously clarified, I consider five different intervals of time of *ESLM* so, in the first stage of our analysis (OLS estimations), three different results of this model will be presented. For the second stage (2SLS) there are also three results to each interval of time in addition to the three different results for each interval of time which will depend on the size of companies on each zone (9 results are expected for this model).

Model 2 :

$$Y = \beta_0 Cons + \beta_1 impro_i + \beta_2 ESLM_{ti} + \beta_3 Pop_i + \beta_4 Area_i + \beta_7 X_i + e \quad (3)$$

Model 2 shows us the dependency of the income factor with respect to $impro_i$, $ESLM_{ti}$, Pop_i , $Area_i$ and X_i . Respecting to available data, these are the variables that best explain the income. Results of this assumption will be found on section 4 (OLS results).

The hypothesis we follow with this model and the other ones is that improvements of public transports on each zone have a bigger impact when it is considered as an *IV* depending on other exogenous variables (see equation 1) than when it just play the role of an ordinary explain variable.

Because Income is explained by *ESLM*, I expect also to have three results for the first stage of the analysis and nine different results on 2SLS analysis; each one with respect to the commuting time and the size of enterprises.

Model 3 :

$$Emplo = \beta_0 Cons + \beta_1 impro_i + \beta_2 Y_i + \beta_3 Pop_i + \beta_4 Area_i + \beta_7 X_i + e \quad (4)$$

On model 3, we advise that the number of employment on each zone of the city depends on the average income of each zone, on the population and the area of that zone and mainly, on the number of stations of Transmilenio on each zone; other control variables were taken in consideration as in previous models.

Furthermore, like in precedents models, the lack of Transmilenio stations in a zone supposed that it does not benefit from public transports enhancements.

Contrary to previous models, employment on a zone i is not explained by *ESLM*. It means that one result is expected on the first stage of our analysis (OLS). Three results are expected on 2SLS models which depend on the size of enterprises.

3 DESCRIPTIVE STATISTICS AND DATA AVAILABLE

Bogotá is the biggest city of Colombia and also one of the most densely populated cities in Latin America. It counts almost eight millions of inhabitants and its density is around 230 people/ha¹³. It is organized following the administrative Parisian model. It means it has 20 sub city urban areas¹⁴ called “Localidades”, each one with its own mayor and an independent budget assigned by the main city hall with respect to the population. Because of their big physical extension, “Localidades” are subdivided in planning zones. In total, there are 112 planning zones in the city. We will develop our research question with respect to each planning zone of the city.

The data we use in our analysis come from various sources.

On 2005, the mobility department of the city decided to make a rather complete study about the mobility behavior in Bogotá¹⁵ with information about

¹³Adapting from Suarez,2005.

¹⁴Urban area is composed by 19 sub city urban areas and one rural. The urban areas count 35.000 hectares. (three times Paris)

¹⁵Secretaria de Movilidad del Distrito; Plan Maestro de Movilidad 2005.

travels of residents. To estimate the job accessibility index, we use the Transport Matrix of Bogotá¹⁶. This matrix encloses information about all possible itineraries “from” and “to” every 112 planning zones of the city. After made some filters of the poll and some estimation, we obtained a summary of the time and the mean distance that people take to get from one zone of the city to any other¹⁷. In parallel, we established the time of travels when people use public transports or their own cars. This information was compiled by the department of mobility of the city.

Table 1: Descriptive statistics for the entire sample and for each variable

Variable	Label	Obs	Mean	Std.	Min	Max
Accessible jobs on 40 minutes	ESLM_40	112	414541,30	226526,60	32151	875709
Accessible jobs on 50 minutes	ESLM_50	112	645827,90	281513,60	84793	1140640
Accessible jobs on 60 minutes	ESLM_60	112	889232,30	287322,00	184320	1320369
Income	Income	112	689235,80	566579,90	196821	3032604
Number of jobs	Employment	112	13997,52	16632,94	197	142052
Improvement degree	Improvment	112	1,92	2,41	0	10
Population on 2007	Population	112	65526,80	50644,21	776	284499
Area of the Zone (hectares)	Area	112	368,13	164,97	86	925
Mean distance of trips	Mean_dist	112	8,69	2,57	5	20
Have at less one car	Car owners	112	6778,81	6065,13	0	27569
Average cost of travel	Aver Cost	112	1878,35	704,86	332	6080
Parks on the zone	Parks	112	48,32	35,19	1	171
Establishments promoting Welfare	Social Welfare	112	63,25	63,83	0	288
Establishments provinding Health	Health	112	3,53	3,24	0	15
Schools	Schools	112	32,76	25,08	0	135
Cultural Establishments	Cultural	112	8,12	8,18	0	40
Establishments promoting Sport and Recreation	Recreation	112	11,42	9,38	0	47
Neighborhoods of each zone	Neighborhoods	112	181,16	105,28	1	526
Working Force	W_Force	112	28543,31	39705,63	0	257862
Micro-enterprises	MiE	112	2584,13	2487,72	12	13826
Small-Medium Enterprises	SME	112	123,19	195,25	0	1674
Large Enterprises	LE	112	26,79	48,35	0	427

Source: Encuesta de Calidad de Vida para Bogotá (ECV) 2007. Secretaria de Planeación del Distrito de Bogotá; Secretaria de Movilidad del Distrito; author calculations.

An emphasis has to be made regarding job accessibility. Index for job accessibility was calculated following the methodology of Prud’Homme, Lee (1999). Actually, as stated beyond in this research, I define “job accessibility” as what Prud’Homme and Lee call “The effective size of labor market”. This theory is based on the assertion that labor market is in function of the commuting time

¹⁶Obtained at the “Secretaria de Movilidad del Distrito” (Department of mobility of the city) and University of Los Andes (Bogota).

¹⁷Olarte, C. (2012) “Heterogeneity of social classes and job accessibility: implications of transport policies in Bogota” Working paper, Centre d’ Economie de la Sorbonne (CES).

from inhabitants homes to zones where people work. With that methodology we are able to estimate the number of jobs they can reach on a specific time. The calculus I made, take into account trips made by inhabitants on private and public transports.

Even if the subject of this study is to analyze the endogeneity and possible causality of enhancement of public transport system on the other variables we proposed, I suppose that private cars also benefit from those improvements. In fact, improvements of public transport system, and more precisely the construction of Transmilenio, is also traduced by an upgrading of corridors alongside Transmilenio corridors. Since the construction of TM, congestion on roads along TM corridors decreased and the average speed had increase. Improvement of public transports traduce by the construction of TM, upgrades travels made by private cars and taxis. Nevertheless, I focus my research only on travels made on public transports.

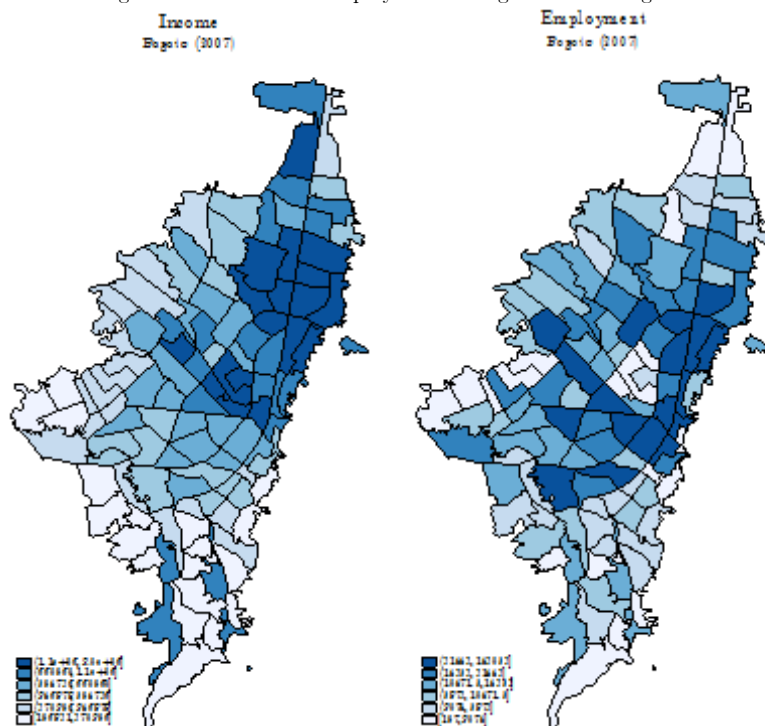
Information of socio-economic indicators such as employment, working force, income, population and area were taken from the census conducted in the city in 2007¹⁸ by the department of planning of the city. We could not have more recent information because the city hall has not made other census for subsequent years.

Data about the number of stations and the number of lines of Transmilenio within or bordering each zone were calculated by the author. To make that calculus, the map of the city with the boundaries of each planning zone and the map and streets where Transmilenio pass were took. Stations within each zone were considered as a station of only this zone. Stations on the border between two or more zones were considered as stations belonging to both zones because both zones benefit of the station.

To complete our model I took in consideration other control variables like the total number of men, the number of neighborhoods on each zone and the number of cars by households. Additionally, variable describing some infrastructure characteristics and amenities of zones were also took in consideration: the number of parks, the number of schools and the number of establishments providing welfare, sport, recreation and cultural events on each zone as well as the number of establishments providing health make also part of the model.

¹⁸Encuesta de Calidad de Vida para Bogotá (ECV) 2007. Secretaria de Planeación del Distrito de Bogotá.

Figure 1: Income and Employment configuration of Bogota



Author calculations based on data from Planning Department
(Secretaria de planeacion del Distrito)

Figure 1 shows a disproportionate configuration of the employment and income in the city. It is easy to identify the axes where jobs and income are concentrated¹⁹. Two different “job centers” are connected on the same two axes. About income, two zones concentrate the highest incomes of the city. These can be distinguished and they are also reliable on the same axis than the two jobs centers. This axis has the particularity that it benefits from the presence of TM. A third job axle is also placed alongside another Transmilenio axle. Regarding the others zones with high levels of income, they are not very far from the TM corridor that goes from the center to the west of the city.

Additionally, jobs are condensed far from boundaries of the city (with the exception of the East border). Even if we see that job centers are placed at the center-east border of the city, this border was always considered as the “center”. Executive and almost the 100% of administrative buildings are place on that zone. Furthermore, is in that place where Bogota was founded and it could not be expanded to the East because of the presence of a big chain of mountains (Los Cerros Orientales). The expansion of the city took place essentially to the south and the north of the city and more recently to the west of the city.

¹⁹Configuration of ESLM depending on travel time can be found on annex 4

Following results from the poll made by the administration on 2005, most of work trips come from the "far" north, the south and the west to these four job centers. Its characteristics are that these are zones with the lowest income in the city with the exception of one zone at the north where people with a big purchasing power have settled few years ago searching to be far from the noise and the pollution of the city.

The "center" of the city condensed more than 50% of jobs in the city and according to the poll, most of the people commuting to this zone of the city come from the north of the city but also from the north-west, the west and the south and the south-west of the city.

Apparently, this configuration of jobs and of income of bogotanians can be directly related to the enhancements of public transport system.

Next section will be focus on the analysis of this possible strong relationship. Results of models proposed on section 2 will give some elements in order to know the existence and the magnitude of this relation.

4 RESULTS AND DISCUSSION

As stated above, an analysis of the correlation matrix and the variance inflation factor (VIF) of all models is recommended.

Correlation matrix²⁰ shows important relationships between some variables²¹, in particular variable "Population" and "Schools". In effect, population has correlations upper than 0.50 with almost all the variables representing the characteristics or the amenities of each zone. These correlations can be interpreted as reasonable. In effect, the higher the population is, the zones will need more schools, parks, establishments of sports and recreation. Likewise, if population is bigger, it can be expectable that the number of neighborhoods will be greater as equal as the area of the zone. In despite of this, correlations just represent some control variables and not make part of the core of our analysis, it let infer that these correlations will not affect the estimations.

Nevertheless, the bigger correlation is between the number of schools in zones and the number of establishments promoting sport and recreation (0.81). It is predictable if we consider that the most of the schools in Bogotá are public. They usually use the infrastructure of the zones (parks, and recreation establishments) to assure recreation and sportive activities of students.

Another important correlation that strikes the goal of this study is between employment and improvements of public transports (number of stations); they

²⁰ Annex 1.

²¹ We consider as important, correlations upper 0.50.

have a correlation index equal to 0.56. It can be also awaited but it seems not to be a problem of multicollinearity if we consider that VIF is smaller than 10. It could also represent an additional argument to suggest that improvements of transports affect directly the number of jobs (model 3).

Other correlations indices upper 0.50 are found. The big indices are principally between amenities variables which is not very surprisingly to find and will not be a source of evils. None of those variables explain each other.

In the other hand, and as indicated previously for some relations between variables, regarding (VIF), no problems have been detected in the models²². On every model it is found that $VIF < 10$ which means that multicollinearity is low and dependent variables are uncorrelated with predictor variables.

The following section will present the results of OLS estimations and 2SLS estimations.

4.1 OLS Results

As noted previously, for the first stage of the analysis, each model depending on the effective size of labor market is regressed three times in accord to the three intervals of times. Table 2 shows results for the three models. There are three different results for models 1 and 2 and one result for model three.

4.1.1 Model 1: Effective Size of Labor Market

Regarding model 1, results show a not negligible R^2 for each travel time which goes for 0.583 when travel time is 40 minutes, to 0.615 when travel time is 60 minutes. Taking into account the analysis of the correlation matrix and also the analysis of the VIF, it can be said that the model explains in an acceptable proportion the dependent variable.

Regarding regressors, table 2 suggests that *improvement* has a positive and big effect on the level of the effective size of labor market. In effect, for an interval of time of 40 minutes, the number of stations of Transmilenio seems to boost the size of labor market of inhabitants of zone i on 15.200 jobs. Furthermore, for intervals of time of 50 and 60 minutes, the presence of improvements of Transmilenio on the zone has a positive relationship with the size of labor market in the order of 20.314 and 21.388 jobs respectively. These three results are significant at the 0.05 level which is not negligible.

Regarding the income on each zone, it is also shown that it has a positive impact on the size of labor market which was predictable. The more the income

²²Annex 2.

is, the higher will be the size of labor market because inhabitants may chose the neighborhood where to live and can be closer to their jobs. Nonetheless, in model 1, income seems not to have a significant impact on *ESLM*.

Other variable that appears to impact significantly the *ESLM* is the mean distance between house of inhabitants and their jobs. This variable has a negative impact on the number of reachable jobs on these intervals of time. For the three different intervals of time, estimators of mean distance between reachable jobs and houses are significant at the 0.01 level. Additionally, its influence looks to be very important on the level of *ESLM*. Actually, an additional kilometer of mean distance reduces the *ESLM* on 5% and 10% for any interval of time. It may seem strange but it can be considered as normal. In effect, concerning the *ESLM*, results confirm the fact that, when individuals live far from their jobs, the number of reachable jobs for their skills in the zone where they live may decrease. It means that the size of labor market of people decreases if their jobs are far from their houses.

Table 2: OLS results for each interval of time

	SLM			Income			Employment
	slm 40	slm 50	slm 60	income 40	income 50	income 60	
Intercept	711.644.313*** (79.704.801)	1045928.176*** (97.313.955)	1313650.550*** (97.115.364)	-136.230.614 (312,658.436)	-134,996.960 (344,254.457)	-209,878.949 (394,996.977)	-3,632.736 (5,667.080)
Improvement	15,220.654** (7,523.774)	20,314.682** (9,185.999)	21,388.184** (9,167.252)	25,161.388 (22,197.682)	26,034.941 (22,362.994)	25,842.513 (22,434.205)	3,274.915*** (534.947)
Job_Accessibility				0.450 (0.292)	0.307 (0.241)	0.301 (0.242)	
Income	0.053 (0.034)	0.054 (0.042)	0.052 (0.042)				0.010*** (0.002)
Population	-0.362 (0.723)	-0.324 (0.883)	-0.441 (0.881)	-3.854* (2.069)	-3.949* (2.074)	-3.918* (2.077)	-0.037 (0.051)
Area	-262.545** (121.573)	-385.040** (148.432)	-388.453** (148.130)	336.586 (360.452)	338.087 (365.529)	337.181 (365.922)	21.047** (8.644)
Mean_dist	-35,218.349*** (7,161.193)	-45,881.804*** (8,743.313)	-50,838.129*** (8,725.470)	6,305.226 (23,276.488)	4,437.753 (23,691.774)	5,676.773 (24,299.335)	-408.308 (509.167)
Car owners	-3.859 (3.465)	-1.873 (4.231)	-0.770 (4.222)	8.649 (10.107)	7.539 (10.101)	7.201 (10.099)	0.229 (0.246)
Aver Cost	59.329** (26.480)	75.797** (32.331)	86.296*** (32.265)	230.366*** (75.460)	235.832*** (75.799)	233.249*** (76.579)	0.221 (1.883)
Parks	-322.012 (713.883)	-781.846 (871.601)	-151.965 (869.823)	4,351.384** (2,031.430)	4,478.577** (2,043.390)	4,287.399** (2,039.829)	-89.009* (50.758)
Social Welfare	-933.344** (385.331)	-1,140.505** (470.462)	-1,007.571** (469.502)	-2,465.939** (1,126.882)	-2,558.896** (1,129.365)	-2,607.019** (1,121.241)	-7.563 (27.397)
Health	-2,964.688 (5,759.372)	-3,710.037 (7,031.788)	-2,941.590 (7,017.438)	-18,035.172 (16,677.106)	-18,382.453 (16,739.421)	-18,646.856 (16,733.046)	234.778 (409.496)
Schools	1,053.693 (1,644.851)	1,486.035 (2,008.247)	1,243.575 (2,004.149)	-4,910.330 (4,769.123)	-4,925.664 (4,790.947)	-4,847.548 (4,789.328)	-32.605 (116.950)
Cultural	-1,027.805 (3,163.854)	85.714 (3,862.843)	-1,576.168 (3,854.960)	17,458.404* (9,036.689)	17,100.412* (9,074.786)	17,612.816* (9,075.872)	984.440*** (224.953)
Recreation	5,871.716* (3,270.352)	5,689.958 (3,992.870)	5,259.589 (3,984.722)	15,048.130 (9,549.043)	16,084.718* (9,510.627)	16,256.850* (9,496.559)	16.487 (232.525)
Neighborhoods	-649.501 (516.125)	-765.556 (630.153)	-1,110.307* (628.867)	638.618 (1,512.266)	583.495 (1,517.503)	683.448 (1,530.195)	-51.926 (36.697)
Obs	112	112.00	112	112	112	112	112
R ²	0.583	0.598	0.615	0.436	0.432	0.432	0.609

Author calculations

Variable *Area* is another variable that seems to be significant from a statistical point of view. In effect, table 2 shows that for any interval of time, it is significant at a 0.05 level. Nevertheless it has always, a negative sign which can be estranged. Actually, regarding the area of the zone and its negative relationship with the size of labor market for each population living on each zone, we can suggest that it could be because jobs and income are concentrated, on a big proportion, on small zones which are placed close to the job centers until the biggest zones are placed at the periphery of the city. People living on the biggest zones are those who are smaller levels of income and are farther from job centers than people living closer to them. In parallel, the bigger the area of each zone is, the higher the number of neighborhood will be and that could be also the reason of the negative sign of neighborhoods parameter.

Estimator for variable Average cost, which represent the average cost of travels that people have to pay to reach their jobs, appear with a positive sign on Table 2. Moreover, this estimator is significant at a 0.05 level with any travel time. It suggests that the more I pay on transport to reach my job, the higher will be the size of my labor market. It could also be expectable if we consider that the price to pay for use of the public improvements we consider in this study (Transmilenio) is higher with respect to the other kind of public transport services excepting taxis. Additionally, all the lines of Transmilenio were constructed on an axis passing by job centers which makes Transmilenio more expensive but at the same time, the quicker transport system to reach job centers.

Regarding amenities of the zones, table 2 shows that four variables seem to have a negative effect on the *ESLM* and two variables have positive impacts. Among these six amenities just one has a significant influence on the dependent variable. *SocialWelfare* has, for the three intervals of time, a negative influence on dependent variable with a significance level of 0.05. It may be due because this variable denotes specially the number of establishments like nursing homes, rehabilitation center, orphanages and establishments promoting the welfare of inhabitants with some problems to interact with society. In that vein, that kind of establishments may not have a positive effect on the labor market because it required a lot of space in the zone and it employs fewer people than other companies with other activates. As well as it occurs with *SocialWelfare*, estimators of variables Parks and establishments promoting cultural activities, *Cultural*, have also a negative incidence on the *ESLM* but do not have significant influence on dependent variable. Reasons may be the same that for *SocialWelfare*. *Parks* and *Cultural* take a lot of vital space of the zone and that avoid the construction of roads, lines of transport systems and jobs centers.

Regarding establishments providing health services, *Health*, results show that it has a negative impact on *ESLM*. It seems to be counterintuitive because hospitals and health centers are supposed to create several jobs. It suggests that the decision to construct hospitals in the city depends on the available area on each zone. In effect, zones where jobs centers are situated do not have as

many available areas to construct hospitals like the peripheral zones have. In addition, the expensiveness of domiciliary public services on jobs centers could also influence this relationship.

In summary, if a comparison of estimators of all variables is made, it can be said that the variable that influences the more the variable *ESLM* is Improvement. In addition, its influence increases with respect to the commuting time. It means that, the bigger is the commuting time of individuals, the bigger will be the influence of improvement of public transports on the effective size of labor market.

In opposition, the other variable that has a significant but negative impact on the *ESLM* with a statistical significance of 0.01, is the mean distance that people have to commute to their jobs. As with *improvement*, "*Mean_dist*" increases with the commuting time, which is expectable. The higher is the mean distance commuted by people the lower is the *ESLM* of inhabitants living on origin zones.

4.1.2 Model 2: Level of Income

Results for model 2 are also divided on three because one of the regressors is *ESLM* and it varies with respect to the commuting time. In the other hand, even if some influences of some estimators are not different from the model 1, there are also five big differences that have to be remarked.

First of all, variable Improvement has a big influence on the level of income of inhabitants of each zone. It suggests that the higher the presence of improvements of public transports is, the greater will be the income of inhabitants. This relationship can be interpreted from two different points of view. In effect, it is not false to suggest that the presence of Transmilenio can boost the income of inhabitants. In the other hand, it can also be said that people with the greater levels of income choose to live on zones where there is more presence of improvements of urban transports. Nevertheless, it is clear that the impact of improvement of urban transports on the level of income of inhabitants cannot be denied. However, even if estimator of improvement is the one that has the bigger influence on the level of income, it also has a problem on OLS results: it is not statistically significant. It represents another additional reason to make a 2SLS analysis on section 4.2.

The other difference between model 1 and model 2 is that the estimators for *Area*, *Mean_Dist*, *Parks*, *Cultural* and *Neighborhoods* are not negative but positive. First of all, regarding the area of the zone, it indicates that the bigger is the zone, the greater is the average income. This relationship is expectable. In effect, as suggested by Anas (1990), Glaeser, Kahn and Rappaport (2000) and other researchers, rich people sometimes prefer to live on suburbs or on zones far

from the city to avoid to be near the noise and the congestion of the job centers. Additionally, is on suburbs or on zones far from the center where rich people find more space to live or to construct bigger houses than those that they find on the downtown or near job centers. In parallel, amenities they prioritized, aside space of houses, are safety, calm (no noise or pollution), green and big spaces to do sport and for their children and proximity to nature in spite of proximity to their jobs or to the center of the city. In that vein, that is the reason to find big and positive influences of estimators for *Parks*, *Recreation*, *Cultural* and *Neighborhoods* where *Parks* and *Cultural* are those that are statistically significant at 0.05 and 0.1 respectively.

Regarding the estimator of *Cultural* and establishments promoting cultural activities, it could be interpreted that cultural manifestations like theatre, opera, concerts among others, are not affordable for people with a limited income. That kind of cultural expressions are revealed to be expensive avoiding the accessibility of people with limited budget and that could be the reason to have a positive and a statistical significance of this estimator on income.

In contrast, *SocialWelfare*, *Health* and *Schools* have a negative and not negligible influence on the level of income but just *SocialWelfare* has is statistical significant (0.05). It can be interpreted as that if there is more hospitals, schools, rehabilitation centers or nursing homes, the congestion and the calm of the zone can be affected and people with bigger income, may not be incentivized to live near that kind of establishments. The easiest and best solution for people with high level of income is to live on far and expensive suburbs with small density and with less number of establishments promoting health, social welfare or education. In opposition, people with small level of income decide to live also far from the center and job centers, but on suburbs or zones with high density and important concentration of hospitals, schools and establishments providing social welfare. Likewise those variables, the negative and statistical significant (0.01) impact of the variable *Population* is explained.

Finally, the other variable that reveals a statistical significance impact on the level of income is *Aver_Cost* which is also expectable because if rich people decide to live far from their jobs, they will be pushed voluntarily to expend more money to commute from their houses to their jobs.

As in model 1, model 2 reveals that the variable that has more influence on dependent variable is the level of improvements of public transports on each zone. Nevertheless, with respect to the level of income, this variable is not significant from a statistical point of view and this leads us to make a 2SLS analysis on section 2.2.

4.1.3 Model 3: Number of jobs

Model 3 shows the influence of some explain variables on the number of jobs on each zone of the city. This model is measured with respect to the same

independent variables than in the model 1 but there is just one result because it does not depend on *ESLM*. Subsequently, even if results are similar, it also has some important differences that should be clarified.

First difference is that income is now statistically significant at 0.01 which it was not the case on model 1. It may suggest that the number of jobs will be bigger on zones where income is higher but it also may be propose that enterprises or companies, settled on zones or near zones where income is higher (Y. Zenou 2000, 2008, S. Ross 1998, J. Kain 1968).

Secondly, table 2 shows that the area of the zone has a positive and a statistical significance (0.05) on the number of jobs on each zone. The bigger is the zone, the more jobs there will be on the zone. This result is the opposite from what was found on model 1, because model 1 tries to determine the impact of the area of the zone on the number of reachable jobs by inhabitants. What was suggested on model 1 was that Area has a negative impact on *ESLM* which is very different if dependent variable is the number of jobs on each zone. *ESLM* takes into account the reachable jobs for each inhabitant with respect to their skills and employment refers about all kind of jobs on each zone. So, this could be the reason of this different relationship on models.

Another regressor that is significant from a statistical point of view is the number of parks on each zone. This variable has a negative influence on the number of jobs on each zone, which is logical because the more of number of parks are on each zone, the less of available space for the settlement of enterprises on the zone will be.

Fourthly, establishments promoting cultural activities seem to boost and to create significantly, the number of jobs on each zone. This result suggests that cultural activities are an important source of employment on the city as equal as establishments providing health. *Schools* seems to have the opposite effect than *Health* or *Cultural*

Aside all other results that are similar and could be read like on model 1, the most important to note is that the variable with the bigger impact on the number of jobs on each zone is, one more time, the number of stations of Transmilenio. In add to this, the estimator of this variable is significant at 0.01 level, which demonstrate the big dependence of employment with respect to the improvements of public transports. It could signify that the number of employment grows up on 3.274 if an additional station of Transmilenio is built on a specific zone. However, it also could denoted that enterprises decide to settled on zones with high level of improvements of public transports and great level of accessibility (R. Fernandez 2008, M. Kawabata 2003).

On the three precedents models table 2 shows that variable *improvement* has always a big and almost the greater influence on dependent variable regardless the commuting time when variable *ESLM* make part of the model. Nonetheless, regarding the statistical significance, this variable is not significant on model 2.

A 2SLS analysis is then necessary to verify if results of these models are not forgetting instruments that can determine or can influence the number of Transmilenio stations on each zone. In other words, a 2SLS will let verify if there exist an endogenous relationship between dependent variables of the three models and the variable *improvement* and therefore, a causal relation from the number of stations on each zone and the dependent variables of each model.

4.2 2SLS Results

As was already defined on section 2.2, the third stage of this analysis is focused on the identification of endogeneity between the level of public transports and dependent variables. As for OLS results, it is essential to clarify that for models 1 and 2 have three results depending on the commuting time.

Additionally, the three models will also be regressed three times according to the definition of the Instrument Variable (*IV*). In effect, it is suggested that the decision to construct a Transmilenio station on a zone depends directly on the density, the working force and the number of enterprises²³ on each zone. In that vein, the size or the number of employees of enterprises based on each zone should be taken into account and should be differentiated. For that reason, this study proposed three different regressions of each model regarding the definition of the *VI* and in accord to the size of companies established on each zone.

4.2.1 Enhancements of PT defined by the number of microenterprises on zone *i*.

Results presented on this subsection correspond to models 1, 2 and 3 when the *VI* is defined by the density, the working force on each zone and the number of microenterprises on each zone.

First of all, table 3 shows that every estimator has the same relationship (sign) and the same statistical significance for model 1 than on OLS results. Improvement is the variable that has the higher positive and statistical influence on the size of labor market on each zone. In parallel, the mean distance between houses and jobs is the variable that has the higher negative and statistical impact on dependent variable. Although this similarity between OLS and 2SLS results for model 1, the impact of the number of stations on the size of labor market is 50% upper with respect to OLS model results. Following this observation, the suggestion that when improvement is treated as an *IV* depending on the number of microenterprises, its impact on the dependent variable increase 50%.

²³ See section 2.2 for definition of what it was called Microenterprises, Middle size enterprises and Large enterprises

Regarding the endogeneity analysis, table 3 reveals that, even if the number of stations on each zone still being significant and greater than in OLS regressions and, even if there is no problem on the choice of instruments that explain IV , variable improvement is not endogenous to the size of labor market. This result is counterintuitive with the objective that this investigation tries to demonstrate but one fact is salvageable and is that, when improvement is considered as a IV , its statistical significance and its impact on $ESLM$ increase on 50%.

Concerning model 2, 2SLS results are not significantly different of those found on OLS estimations. Similarly, relationships between variables are respected and every estimator preserves the same sign, the same statistical significance and almost the same level. It means that the number of stations on each zone is not yet statistically significant but its effect on $ESLM$ is almost 70% upper than in the OLS models.

Table 3: 2SLS results for each interval of time when:

$$impro = \pi_0 + \pi_1 WF_i + \pi_2 MiE_{si}$$

	SLM			Income			Employment
	40 minutes	50 minutes	60 minutes	40 minutes	50 minutes	60 minutes	
Intercept	703,963.433*** (76,209.254)	1031718.345*** (93,583.480)	1302852.910*** (93,355.711)	-121,333.320 (314,339.274)	-108,458.695 (323,163.063)	-159,857.237 (372,672.907)	-10,059.739 (6,948.676)
Improvement	22,117.543* (12,173.235)	32,521.548** (14,948.495)	31,071.155** (14,912.112)	40,792.818 (38,949.393)	41,285.053 (36,786.388)	43,053.538 (36,908.604)	7,588.386*** (1,109.942)
SLM				0.394 (0.306)	0.258 (0.236)	0.243 (0.238)	
Income	0.047 (0.033)	0.043 (0.041)	0.044 (0.040)				0.007** (0.003)
Population	-0.441 (0.671)	-0.423 (0.824)	-0.578 (0.822)	-3.634* (2.060)	-3.729* (1.922)	-3.694* (1.925)	-0.012 (0.061)
Area	-0.020** (0.010)	-0.029** (0.012)	-0.028** (0.012)	0.025 (0.031)	0.024 (0.029)	0.024 (0.029)	0.002** (0.001)
Mean_dist	-36,208.470*** (6,614.546)	-47,168.293*** (8,122.534)	-52,548.061*** (8,102.765)	6,706.096 (23,290.065)	4,535.453 (22,132.616)	5,150.545 (22,823.788)	-177.267 (603.107)
Car owners	-3.272 (3.307)	-0.892 (4.061)	0.086 (4.052)	9.093 (10.239)	8.080 (9.555)	7.885 (9.574)	0.484 (0.302)
Aver Cost	59.490** (24.888)	76.436** (30.562)	86.289*** (30.487)	236.064*** (75.742)	241.762*** (70.807)	240.669*** (71.623)	0.977 (2.269)
Parks	-324.728 (671.884)	-774.181 (825.060)	-159.629 (823.052)	4,440.929** (2,038.148)	4,544.679** (1,907.578)	4,388.467** (1,905.703)	-62.548 (61.262)
Social Welfare	-905.372** (362.732)	-1,094.140** (445.428)	-966.178** (444.344)	-2,472.263** (1,130.620)	-2,569.331** (1,053.969)	-2,614.283** (1,047.196)	4.153 (33.074)
Health	-3,834.870 (5,490.552)	-5,175.738 (6,742.291)	-4,215.141 (6,725.881)	-19,269.376 (17,005.259)	-19,584.524 (15,888.868)	-19,986.315 (15,885.761)	-174.459 (500.622)
Schools	763.630 (1,588.269)	979.882 (1,950.364)	839.208 (1,945.617)	-5,357.926 (4,889.940)	-5,340.447 (4,567.202)	-5,325.691 (4,571.254)	-191.838 (144.817)
Cultural	-353.538 (3,065.944)	1,227.071 (3,764.920)	-595.260 (3,755.757)	18,263.733* (9,293.874)	17,930.487** (8,707.936)	18,462.362** (8,684.017)	1,309.516*** (279.549)
Recreation	5,340.337* (3,202.302)	4,722.659 (3,932.365)	4,513.505 (3,922.794)	13,805.633 (9,940.897)	14,835.026 (9,245.787)	14,872.610 (9,239.208)	-385.920 (291.982)
Neighborhoods	-699.074 (482.502)	-834.504 (592.503)	-1,196.246** (591.061)	660.919 (1,516.013)	603.795 (1,415.777)	677.960 (1,432.348)	-52.580 (43.994)
Endogeneity - Wu - Hausman F Test	0.44933	0.94049	0.58849	0.22424	0.20373	0.25837	48.18274***
Sargan	3.337	4.646*	1.956	4.002	4.158	4.459	2.690
Heteroscedasticity	19.211	18.153	15.928	36.676 ***	36.388 ***	34.986 ***	19.623
R ²	0.576	0.586	0.605	0.432	0.428	0.427	0.346

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Author calculations

As in the previous model, selected instruments does not have problems of over identification but, the analysis of endogeneity displays that “improvement” is not endogenous to the income of inhabitants of each zone. It was not what the study had expected and it will be explained at the end of the reading of all 2SLS results.

In the other hand, results for model 3 are slightly different. Variables like *income*, *Population*, *Aver_Cost* and *Car_owners* preserve the same impact and the same statistical inference on the dependent variable than on OLS results. Likewise, the mean distance between homes and jobs has the same relationship than in OLS model but its influence on independent variable increased on 100%. In addition, even if the variable “*Area*” has the same statistical influence (0.05) on the number of jobs than in OLS model, its “real” impact is marginally (0.002) contrary to the number of parks on each zone that is no longer statistical significant but its real influence on employment is also negligible. The other group of variables does not have a significant influence on the number of jobs on each zone except two variables that are very significant and that have a notably influence on the number of jobs.

The two regressors that have a significant effect on the employment on each zone are the number of establishments promoting cultural activities and the number of stations on each zone. Regarding *Cultural*, its statistical significance is high (0.01). This result could signify that the implantation of one additional establishment promoting cultural activities could raise on 1.309 the number of jobs on each zone. In the other hand, this result could also be interpreted differently; the target population of that kind of establishments could be employees but this study does not have more enough information to make this statement.

Respecting the number of stations, like in OLS results, it can be observed that this is the regressor with the bigger effect on the explained variable. When *improvement* is considered as endogenous and when it depends on the number of microenterprises, its influence on the number of jobs on each zone is the double than on OLS model. In addition its statistical significance still been very high (0.01). In effect, an additional station on each zone supposes to boost the number of jobs on 7,560. Besides, it can also be supposed that stations of Transmilenio were built looking to be close to the employments centers. Both hypotheses do not move away from the goal of this analysis, which is to demonstrate the causal relation between the number of jobs on each zone and the level of improvements of public transports.

Regarding this statement, endogeneity test reveals that *improvement* is endogenous to the number of jobs on each zone. In other words, the number of stations of Transmilenio is endogenous with respect to the number of jobs. This result implies that the hypothesis is verified which brings to prove the following hypothesis:

$$H_0 : \rho(\textit{improvement}) = 0(\textit{exogeneity})$$

$$H_1 : \rho(\textit{improvement}) = 0(\textit{endogeneity})$$

The p-value of Wu-Hausman F test is statistically significant (0.01) which leads to reject null hypothesis of exogeneity, H_0 . The endogeneity in the relation between *improvement* and the number of jobs on each zone is demonstrated.

This result is even truer if, the test of over identifying restrictions, “Sargan N*R-sq test”²⁴, and heteroscedasticity test (Pagan-Hall²⁵) tests are taken into consideration. In fact, those tests demonstrate that instruments have no problems of over identification and heteroscedasticity (no rejection of null hypothesis). It means that instruments, are exogenous and the residuals of the main model are uncorrelated with the set of instrumental variables. Instruments were well chosen.

Considering results for models 1 and 2, OLS results appear to be consistent. In effect, *improvement* is not endogenous when it is regressed depending on density, working force and microenterprises. By contrast, OLS results are not consistent for model 3. In fact, suspicions of endogeneity of *improvement* are corroborated which represent an undeniable causality relation from enhancements of urban transports on the number of jobs when the number of stations are defined by microenterprises.

4.2.2 Improvement regressed with respect to “small and middle size” enterprises

Results presented on this subsection correspond to models 1, 2 and 3 when the variable suspected to be endogenous is defined by the density, the working force on each zone and the number of middle size enterprises on each zone.

Results for model 1 express exactly the same relationships than on precedent subsection. This means that, there is no evidence to reject the use of instrumental variables on the model. In effect, Sargan test shows that there is no over identification of instrumental variables likewise with the heteroscedasticity test that confirm that the variables are homoscedastic. Nevertheless, as on precedent model, Endogeneity – Wu – Hausman F test allows to accept null hypothesis of exogeneity; the number of stations on each zone of the city are not endogenous on the effective size of labor market. Taking into account those precedent observations, it can be confirmed that OLS results are consistent for model 1.

²⁴The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Under the null, the test statistic is distributed as chi-squared in the number of (L-K) overidentifying restrictions. A rejection casts doubt on the validity of the instruments.

²⁵ H_0 : Disturbance is homoscedastic

In contrast, results for model 2 suppose different interpretations. Variable *Population* continue to have the same impact on dependent variable but it ceases to have a statistical significance as well as *Social_Welfare* preserve its influence on Income but it is no more significant at 0.05 but at 0.1. In the other hand, establishments providing health services increase their negative influence on the level of income on 70% more than when *improvement* is defined by *SME*. Additionally, this variable has a statistical significance of 0.1 which is not very representative but that suggest that if the number of middle size enterprises is taken into account at the moment to define variable improvement, an additional establishment providing health in the zone will lower the income of inhabitants on 34.500 COP²⁶. It represents a decrease of 8%²⁷ of the minimum wage established by the government on 2007. These results could suggest that inhabitants with higher incomes prefer to live on zones with few promoters entities of health. The difference of income of inhabitants living on zones with several promoter entities of health with respect to income of inhabitants that prefer to live on zones with few number of that kind of establishments is 34.500 COP lower, which represents 8% of minimum wage for that year.

Regarding the number of stations of Transmilenio on a zone and its effect on the level of income, table 4 reveals that its influence is notably. In effect, results suggest that one additional station of Transmilenio will boost the income of inhabitants by 197.000 COP²⁸ which represents 28% of the average wage of citizen and 45% of minimum wage decreed by Colombian government on 2007²⁹. But this result could also be interpreted as that, people that live on a zone with an additional station of Transmilenio with respect to another zone, will have an income on 197.000 COP higher and this with a statistical significance of 0.01. In addition, results show that variable *improvement* is endogenous and that there is no evidence against the veracity of instruments used on the model. On the same way, instruments do not present problems of heteroscedasticity.

Precedent results affirm the hypothesis that, when *improvement* depends on the density of the zone, the working force of the zone and the number of “middle size” enterprises, there is a causal relationship of the enhancements of public transports on the level of income of inhabitants.

²⁶ COP: Colombian Pesos. The exchange rate on 2007 was: 1 USD = 2078 COP; 34.500 COP = 16.6 USD

²⁷The minimum wage represents the minimum wage that a worker have to be paid monthly; on 2007, the minimum wage in Colombia was 433.700 COP = 208 USD.

²⁸ 197.000 COP = 95 USD on 2007.

²⁹ See table 1: Average wage per month for 2007 at the city Bogota was 687.235 COP which on dollars of 2007 is 331 USD.

Table 4: 2SLS results for each interval of time when:

$$impro = \pi_0 + \pi_1 WF_i + \pi_2 SME_{si} + u$$

	SLM			Income			Employment
	40 minutes	50 minutes	60 minutes	40 minutes	50 minutes	60 minutes	
Intercept	708,406.951*** (76,485.215)	1032649.269*** (94,123.698)	1295251.120*** (94,719.511)	-39,144.142 (367,614.009)	64,556.348 (407,977.334)	139,053.591 (472,640.016)	-15,271.339 (9,751.496)
Improvment	19,046.838	31,878.230*	36,324.391**	193,049.558*** (51,658.167)	196,398.682*** (52,688.143)	197,621.601*** (53,398.809)	11,189.875***
SLM	(13,579.695)	(16,711.348)	(16,817.133)				(1,731.346)
Income	0.049 (0.033)	0.044 (0.041)	0.040 (0.041)	-0.042 (0.364)	-0.129 (0.304)		0.005 (0.004)
Population	-0.453 (0.670)	-0.425 (0.824)	-0.558 (0.829)	-2.948 (2.410)	-2.999 (2.423)	-3.039 (2.430)	0.002 (0.085)
Area	-0.020** (0.010)	-0.030** (0.012)	-0.028** (0.012)	0.028 (0.036)	0.025 (0.037)	0.025 (0.037)	0.002* (0.001)
Mean_dist	-36,304.776*** (6,597.005)	-47,188.469*** (8,118.360)	-52,383.304*** (8,169.750)	-3,876.159 (27,272.652)	-8,445.628 (27,954.481)	-10,747.558 (28,905.269)	-64.314 (841.086)
Car owners	-3.463 (3.319)	-0.932 (4.084)	0.413 (4.110)	16.128 (12.020)	16.175 (12.106)	16.338 (12.157)	0.707* (0.423)
Aver Cost	59.005** (24.830)	76.334** (30.556)	87.119*** (30.750)	263.892*** (88.634)	272.159*** (89.315)	276.431*** (90.516)	1.546 (3.166)
Parks	-341.290 (670.627)	-777.651 (825.282)	-131.297 (830.506)	4,736.085** (2,382.514)	4,665.679* (2,402.630)	4,746.860** (2,403.754)	-43.124 (85.502)
Social Welfare	-914.176** (362.040)	-1,095.984** (445.532)	-951.116** (448.352)	-2,214.583* (1,322.035)	-2,327.799* (1,328.025)	-2,343.030* (1,321.256)	14.479 (46.158)
Health	-3,532.719 (5,506.433)	-5,112.437 (6,776.288)	-4,732.048 (6,819.183)	-33,668.621* (20,008.288)	-34,241.089* (20,151.585)	-34,309.575* (20,177.473)	-528.838 (702.044)
Schools	878.794 (1,599.799)	1,004.009 (1,968.733)	642.191 (1,981.195)	-10,110.038* (5,765.620)	-10,031.781* (5,802.116)	-10,044.785* (5,819.265)	-326.908 (203.967)
Cultural	-592.708 (3,093.157)	1,176.964 (3,806.479)	-186.098 (3,830.575)	28,044.848** (10,974.733)	28,278.183** (11,094.413)	28,076.871** (11,070.517)	1,590.028*** (394.363)
Recreation	5,621.119* (3,240.725)	4,781.483 (3,988.079)	4,033.156 (4,013.324)	1,418.761 (11,786.867)	1,867.156 (11,832.242)	1,960.132 (11,853.673)	-715.235* (413.177)
Neighborhoods	-696.189 (481.055)	-833.899 (591.992)	-1,201.182** (595.740)	186.329 (1,773.443)	110.003 (1,784.926)	26.772 (1,809.478)	-55.965 (61.332)
Endogeneity - Wu - Hausman F Test	0.10503	0.61123	0.99825	23.88448***	23.52866 ***	23.15102***	432.31390 ***
Sargan	4.229	5.052*	1.694	0.898	0.907	0.973	0.496
Heterocedasticity	19.863	18.270	15.567	17.900	17.902	17.916	9.486
R ²	0.579	0.5869	0.598	0.104	0.093	0.089	-0.270

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Author calculations

Regarding the variable *Mean_Distance* and comparing its estimator with the one when instrumental variable was the number of microenterprises on each zone, table 4 reveals a change of sign. In effect, there is a negative influence of the mean distance on the level of income and, even if is not significant from a statistical point of view its influence is bigger than on precedent models. It suggests that if the distance of between homes and jobs increase, it has a negative impact on income of inhabitants. In other words, what this result suggests is that *SME* are not placed close to zones where inhabitants have a small income or that *SME* are situated close to people that have higher incomes. In the other hand it could also reveal that people with greater incomes can decide to live closer to *SME* than people with modest incomes, thus, the negative influence of *Mean_Distance* on *Income*.

Concerning the variable *Schools*, it is also seen that its influence on the dependent variable is doubled with respect to the previous model and their relationship remains negative. In addition, this variable is statistically significant at 0.1 which convert it on a not negligible variable on the model.

Finally, the two variables that have the biggest positive influence on the level of income are the number of establishments promoting cultural events and the number of stations of Transmilenio. Regarding *Cultural*, its influence on the level of income increases on 60%. It means that, if an additional establishment promoting cultural activities is built on a zone, this could boost the income of inhabitants by 28.000 COP and this with a statistical significance of 0.05.

Results for model 3 do not differ than the model with “microenterprises” as an instrumental variable. Actually, all the variables have the same influence on the income of inhabitants. Statistical significance remains the same as on previous model. The variables that most significantly affect the level of income are *Cultural* and *Improvement*. Regarding *Cultural*, it is displayed that its estimator increased by 20% over the precedent model where instrumental variable was *MiE*. In the same way, the effect of the number of Transmilenio stations on the level of income of inhabitants increased on 50%.

Finally, the endogeneity test gives as results the rejection of null hypothesis, which implies that *improvement* is not exogenous and that 2SLS results are more consistent than those found on OLS model. In parallel, the Sargan test and the heteroscedasticity test validate the use of instrumental variable on the model.

Causality of improvement on the level of income is demonstrated.

4.2.3 Improvement regressed with respect to “Large” enterprises

This final subsection corresponds to results of the three models when the variable *improvement* depends on the density, the working force on each zone and the number of enterprises with more than 50 employees.

As on previous subsections, results for model 1 do not change considerably. The signs, the level of the impact on dependent variable and the statistical significance of estimators remain unchanged. By the side of instrumental variables, the Sargan test and the heteroscedasticity test show that there is no evidence to reject the use of them and that there are no problems of heteroscedasticity. Nonetheless, the endogeneity test determines that there is no a causality relationship between the number of stations of Transmilenio and the effective size of labor market. It implies that OLS results are consistent. A 2SLS analysis is not verified and the use of instrumental variables is not justified. Endogeneity was no proved.

Table 5: 2SLS results for each interval of time when:

$$impro = \pi_0 + \pi_1 WF_i + \pi_2 LE_{si} + u$$

	SLM			Income			Employment
	slm 40	slm 50	slm 60	income 40	income 50	income 60	
Intercept	696,094.362*** (78,433.247)	1012411.954*** (98,263.896)	1273635.499*** (99,657.567)	-38,665.391 (368,709.311)	66,763.779 (411,342.154)	142,725.122 (478,936.966)	-17,930.795 (11,508.759)
Improvment	27,555.484* (16,193.431)	45,863.278** (20,287.693)	51,261.920** (20,575.433)	193,936.451*** (60,240.124)	198,377.717*** (62,034.597)	199,520.165*** (63,074.685)	13,027.699*** (2,376.114)
SLM				-0.045 (0.376)	-0.134 (0.316)	-0.164 (0.320)	
Income	0.043 (0.034)	0.034 (0.043)	0.030 (0.043)				0.003 (0.005)
Population	-0.420 (0.679)	-0.371 (0.850)	-0.501 (0.863)	-2.944 (2.419)	-2.989 (2.439)	-3.031 (2.445)	0.009 (0.100)
Area	-0.019* (0.010)	-0.028** (0.013)	-0.027** (0.013)	0.028 (0.036)	0.025 (0.037)	0.025 (0.037)	0.003* (0.001)
Mean_dist	-36,037.920*** (6,685.674)	-46,749.858*** (8,376.045)	-51,914.820*** (8,494.842)	-3,937.800 (27,409.879)	-8,611.248 (28,206.642)	-10,942.835 (29,226.119)	-6.675 (981.010)
Car owners	-2.935 (3.402)	-0.064 (4.263)	1.340 (4.323)	16.169 (12.127)	16.279 (12.275)	16.441 (12.342)	0.821* (0.499)
Aver Cost	60.349** (25.180)	78.544** (31.546)	89.479*** (31.994)	264.054*** (88.986)	272.547*** (89.928)	276.870*** (91.218)	1.836 (3.695)
Parks	-295.399 (680.658)	-702.224 (852.752)	-50.733 (864.846)	4,737.804** (2,387.926)	4,667.223* (2,413.226)	4,751.262** (2,415.003)	-33.212 (99.875)
Social Welfare	-889.780** (367.430)	-1,055.887** (460.329)	-908.288* (466.858)	-2,213.082* (1,325.650)	-2,324.718* (1,334.761)	-2,339.698* (1,328.044)	19.748 (53.914)
Health	-4,369.952 (5,639.089)	-6,488.536 (7,064.847)	-6,201.869 (7,165.047)	-33,752.496* (20,258.235)	-34,428.086* (20,469.202)	-34,485.508* (20,494.024)	-709.676 (827.441)
Schools	559.686 (1,651.439)	479.516 (2,068.981)	81.976 (2,098.325)	-10,137.719* (5,856.446)	-10,091.636* (5,909.067)	-10,102.750* (5,930.858)	-395.833 (242.321)
Cultural	70.010 (3,202.248)	2,266.227 (4,011.887)	977.351 (4,068.788)	28,101.822** (11,173.038)	28,410.205** (11,350.086)	28,194.966** (11,307.016)	1,733.172*** (469.876)
Recreation	4,843.099 (3,373.519)	3,502.709 (4,226.461)	2,667.288 (4,286.405)	1,346.608 (12,073.174)	1,701.704 (12,188.071)	1,801.528 (12,222.204)	-883.283* (495.007)
Neighborhoods	-704.184 (487.194)	-847.041 (610.374)	-1,215.219** (619.031)	183.564 (1,779.510)	103.703 (1,795.660)	18.774 (1,822.401)	-57.692 (71.488)
Endogeneity - Wu - Hausman F Test	0.66310	1.91267	2.60483	15.17193 ***	14.91195***	14.56189 ***	417.58137***
Sargan	3.166	3.391	1.342	0.985	1.001	1.104	1.553
Heteroedasticity	19.129	16.858	13.779	16.528	16.328	16.424	7.043
R ²	0.568	0.561	0.566	0.101	0.085	0.081	-0.725

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Author calculations

Concerning results for model 2, table 5 reveals the same conclusions than when variable suspected to be endogenous is regressed with respect to the number of middle-size enterprises. This denotes that the variable *Health* is the variable that has the bigger negative impact on the level of income of inhabitants. As it was explained on precedent subsection, it could imply that inhabitants with higher incomes chose to live on zones with few establishments that provide health. The average income of people living in areas with one less entity that provides health services with respect to the average income of people living in areas with an additional facility of this kind, is superior by 34.500 COP (8% of the minimum wage).

Concerning the impact of the number of stations of Transmilenio on the income of inhabitants, results reveal that first of all, this variable is, from far, the variable that impacts considerably the income of inhabitants. Like on precedent subsection, it signifies that income of inhabitants with an additional station on

a zone, income of will be in average, 197.000 COP (95 USD) higher with an additional station of TM.

In effect, concerning the endogeneity test, table 5 reveals that the number of station of TM is endogenous. It suggests that, a causal relationship of the variable *improvement* on the level of income is verified when endogenous variable depends on the number of large enterprises, *LE*, on each zone.

In the same way, results for model 3 suggest the same interpretations. In fact, variable "employment" also depends considerably on the variable *improvement*. An additional station of TM on a zone will boost by 13.027 the number of jobs on this zone. As on model 2, the level of enhancements of public transport is endogenous which means that there is causal relationship of the number of stations of TM on the number of jobs on each zone. For both models, there is no evidence against the use of the instruments and there are no problems of heteroscedasticity of them.

Causal relation is confirmed on model 2 and 3 when the number of "large" enterprises plays the role of instrumental variable.

Table 6: Summary of causality test results

	Test	SLM	Income	Employment
Micro	Endogeneity	No	No	Yes
	Sargan	No	No	No
	Heteroscedasticity	No	Yes	No
PME	Endogeneity	No	Yes	Yes
	Sargan	No	No	No
	Heteroscedasticity	No	No	No
Large	Endogeneity	No	Yes	Yes
	Sargan	No	No	No
	Heteroscedasticity	No	No	No

Author calculations

5 CONCLUSIONS

From the beginning the study focuses on the fact that public transports improvements affect job accessibility, income of inhabitants and the number of jobs on each zone of the city. In the same way, it was suggested that enhancements of public transports could have a causal relation with income, employment and job accessibility and that was the reason to analyze the possible endogeneity of this variable. The proxy for improvements of public transports was defined by the number of stations of TM on each zone. This variable which is suspected to be the endogenous one, is defined by some instruments on each model that their use appear to be accepted by different test (Sargan test Heteroscedasticity test).

Despite the lack of data for several periods, this “cross section” analysis displays interesting results that suggest some particular conclusions that have to be enounced.

Compilation of results reveals that, on OLS regressions for the effective size of labor market, improvements of public transports have always a direct and an important effect on dependent variables with a statistical significant level. It corroborates findings of precedent studies (Ihlanfeldt, Sjoquist; 1991).

In fact, this variable was always the one that had the biggest positive impact on the three different models regardless the commuting time but its amplitude appears to be lower than expected and in some cases, its statistical significance may not reveal its real importance (model 2).

Those results motivate the use of the 2SLS method that explains the suggestion of instrumental variables in order to capture the real and thus suspected impact of the number of stations of Transmilenio on dependent variables.

Regarding the size of labor market, variables that had the bigger but a negative effect on the “*ESLM*” were the mean distance traveled and the area of each zone which is consistent with Prud’Homme and Lee (1999) with respect to the “three S” (Size, Speed and Sprawl). Endogeneity of improvements of public transports on the *ESLM* was not proved at all. In effect, regardless the size of companies on each zone and the working force of each zone (instrumental variables), the number of stations of TM on each zone has not a causal effect on the number of reachable jobs on 40, 50 and 60 minutes. This may imply that, for the case of Bogotá, the decision to enhance public transports at some zones of the city, or to be more precisely, the decision to construct a TM station on a zone, could not be motivated by the potentially reachable jobs from these zones. The *ESLM* depends “significantly”, on the number of Transmilenio station but there is no evidence to assert that the degree of enhancements of public transport has a causality effect on the size of labor market to inhabitants.

In the other hand, as OLS results for model 2 reveal, the number of stations of Transmilenio on each zone of the city has not a statistical influence on income. Nevertheless, this study supposes that this variable should have a greater statistical influence on the income. In effect, after applying 2SLS method and excepting when the endogenous variable is defined by the number of micro enterprises (instrumental variable), results show that improvements of public transports are significant and have an important impact on the income of inhabitants.

In addition, when instrumental variables are the working force on each zone and the number of small and middle size enterprises or the number of large enterprises, enhancements of public transports are endogenous to the income thus a causality relation of improvements of public transport on the income is demonstrated. Furthermore, the use of instrumental variable reveals the underestimation of the effect of the number of stations of Transmilenio on the income of inhabitants.

These results allow to do some suggestions: First of all it might be said that people with highest incomes are those that can settle on zones with a high number of Transmilenio stations. Secondly, in order to understand results, microenterprises appear to be established on low income zones contrary to small, medium and large enterprises that seem to settle on zones with high level of income. Those conclusions suggest that small, medium and large enterprises seem to be in a bigger capacity, like people with high incomes, to settle on zones benefiting more from enhancements of public transports. Until 2007, zones that have more Transmilenio stations appear to attract people and enterprises with the highest incomes of the city.

In relation to employment, this study showed that improvements of public transports are always statistically significant. Moreover, its influence increases radically when the number of jobs on each zone depends on instrumental variables. Likewise, this is even more evident if the size of enterprises increases. Like with the income, the number of Transmilenio stations is endogenous to the number of jobs hence, its causality relationship is verified no matter the size of the companies. In a first time, results suggest that public transports improvements incentive the creation of jobs in a zone. Nevertheless, it could also suggest that jobs (and not workers) decide to settle on zones with better level of public transport improvements. As denoted previously, biggest enterprises, which are those that create almost 70% of employments in the city, prefer to be close to Transmilenio stations. Microenterprises seem not to be capable to benefit from public transports enhancements.

Results shed light on the causal effect of improvements of public transport system on the income and the number of jobs on each zone of the city of Bogotá. In particular, they indicate that the more is the presence of stations of Transmilenio the bigger will be the income and the number of jobs. Moreover, this causal relationship is even greater if the size of companies that settle on those zones is bigger. For now, zones where the largest enterprises are located, benefit from urban transportation enhancements. After testing this causality, this analysis might suggest to build and expand the Transmilenio system to zones where microenterprises, which assure 30% of formal jobs in the city, are established.

6 Annexes

Annex 1 Average commuting time according to type of transport

Conveyance	Commuting time (min)
By walk	27
Bicycle	31
Other	31
Taxi	34
Moto	35
Private cars	40
Transmilenio	41
Truck	46
Intermunicipal bus	46
Scolar bus	48
Private bus	49
Public transport diferent of TM	56
Average	40

Author calculations

Annex 2: Correlation Matrix

	1	2	3	4	5	6	7	
	ESLM_60	Income	Employment	Improvement	Population	Area	Mean_dist	
1	ESLM_60	1,00						
2	Income	0,31	1,00					
3	Employment	0,36	0,44	1,00				
4	Improvement	0,38	0,20	0,56	1,00			
5	Population	-0,24	-0,28	-0,03	0,09	1,00		
6	Area	-0,43	-0,01	0,05	-0,01	0,48	1,00	
7	Mean_dist	-0,41	0,13	-0,07	-0,16	-0,34	0,19	1,00
8	Car owners	0,04	0,00	0,06	0,14	0,57	0,24	-0,27
9	Aver Cost	0,21	0,47	0,17	0,05	-0,31	0,04	0,35
10	Parks	-0,21	-0,05	-0,03	0,06	0,74	0,43	-0,25
11	Social Welfare	-0,30	-0,38	-0,03	-0,01	0,65	0,21	-0,26
12	Health	0,00	-0,17	0,20	0,27	0,45	0,17	-0,29
13	Schools	0,08	-0,03	0,33	0,36	0,70	0,27	-0,32
14	Cultural	-0,11	-0,07	0,34	0,05	0,45	0,12	-0,21
15	Recreation	0,20	0,08	0,32	0,46	0,56	0,21	-0,32
16	Neighborhoods	-0,32	-0,02	-0,03	0,02	0,60	0,53	-0,06

	8	9	10	11	12	13	14	15	16
	Car owners	Aver Cost	Parks	Social Welfare	Health	Schools	Cultural	Recreation	Neighborhoods
1									
2									
3									
4									
5									
6									
7									
8	1,00								
9	-0,02	1,00							
10	0,45	-0,23	1,00						
11	0,24	-0,36	0,49	1,00					
12	0,22	-0,22	0,33	0,40	1,00				
13	0,55	-0,11	0,53	0,46	0,53	1,00			
14	0,13	-0,20	0,39	0,54	0,38	0,64	1,00		
15	0,51	-0,06	0,47	0,43	0,49	0,81	0,38	1,00	
16	0,37	-0,06	0,63	0,30	0,22	0,42	0,32	0,30	1,00

Author calculations

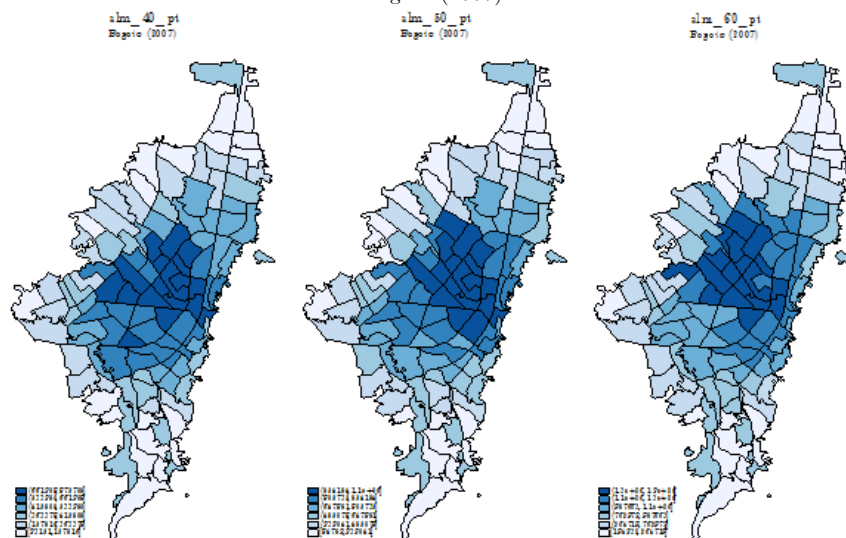
Annex 3: VIF calculations

SLM			Income			Employment		
Variable	VIF	1/VIF	Variable	VIF	1/VIF	Variable	VIF	1/VIF
education	7.72	0.129463	education	7.71	0.129664	education	7.72	0.129667
pop_2007	6.11	0.163555	pop_2007	6.98	0.143186	pop_2007	6.08	0.164413
rec_sport	4.27	0.234157	rec_sport	4.23	0.236516	rec_sport	4.27	0.234169
culture	3.64	0.274602	culture	3.69	0.270689	culture	3.04	0.328818
parla	2.95	0.338479	parla	2.76	0.361931	parla	2.86	0.349210
soe_welf	2.75	0.364112	soe_welf	2.74	0.363410	soe_welf	2.74	0.364398
employment	2.56	0.390699	slm_60_pt	2.70	0.370087	number_neig	2.11	0.474836
number_neigh	2.15	0.465232	employment	2.31	0.432448	have_a_car	2.00	0.498964
Improvement	2.07	0.482405	number_neigh	2.18	0.457705	superficie	1.82	0.547994
income	2.02	0.494263	superficie	2.12	0.470744	income	1.73	0.577608
have_a_car	2.02	0.494576	mean_dist	2.08	0.481033	aver_cost_~1	1.58	0.632700
superficie	1.94	0.516430	Improvement	2.08	0.481326	health	1.58	0.633736
health	1.58	0.631596	have_a_car	2.02	0.494343	mean_dist	1.54	0.649947
aver_cost_~1	1.58	0.632610	health	1.56	0.640362	Improvement	1.50	0.666793
mean_dist	1.55	0.645866	aver_cost_~1	1.56	0.642255			
Mean VIF	3.00		Mean VIF	3.05		Mean VIF	2.90	

Author calculations

Annex 4: Size of Labor Market respecting the zone where people live (UPZ)

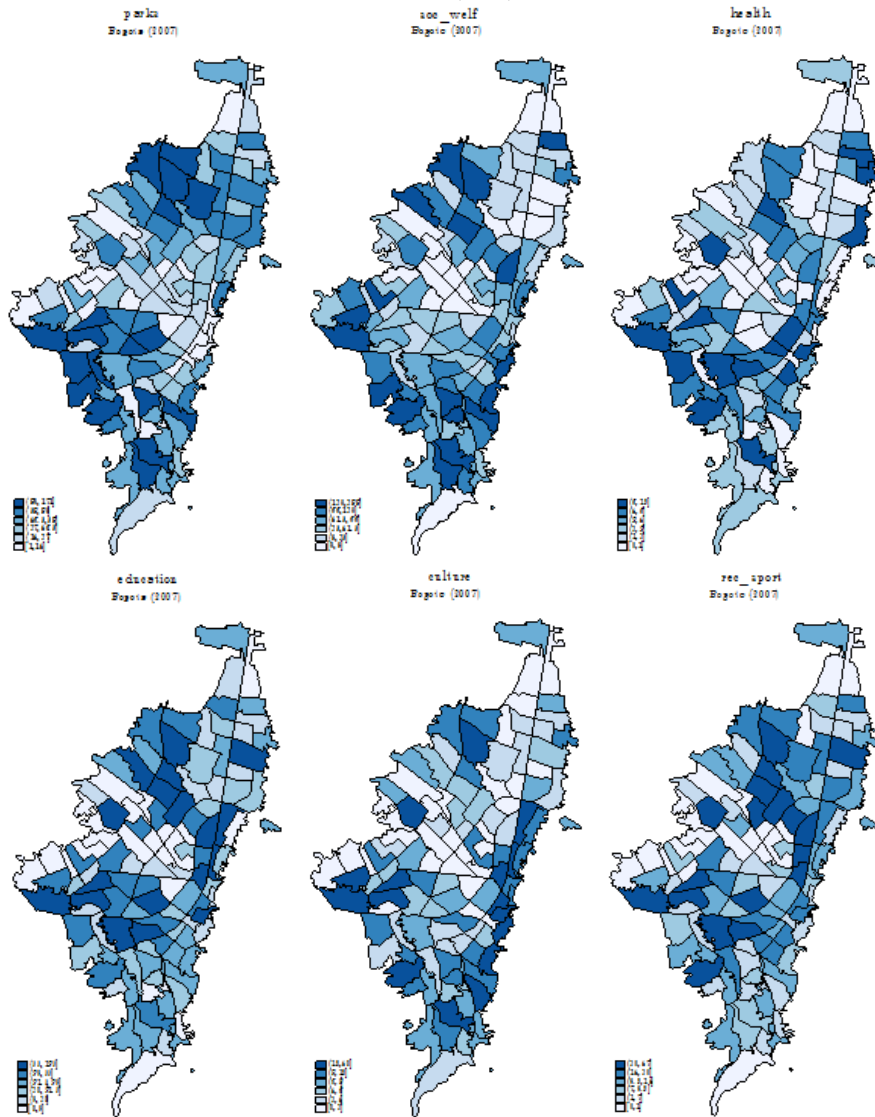
Bogota (2007)



Author calculations

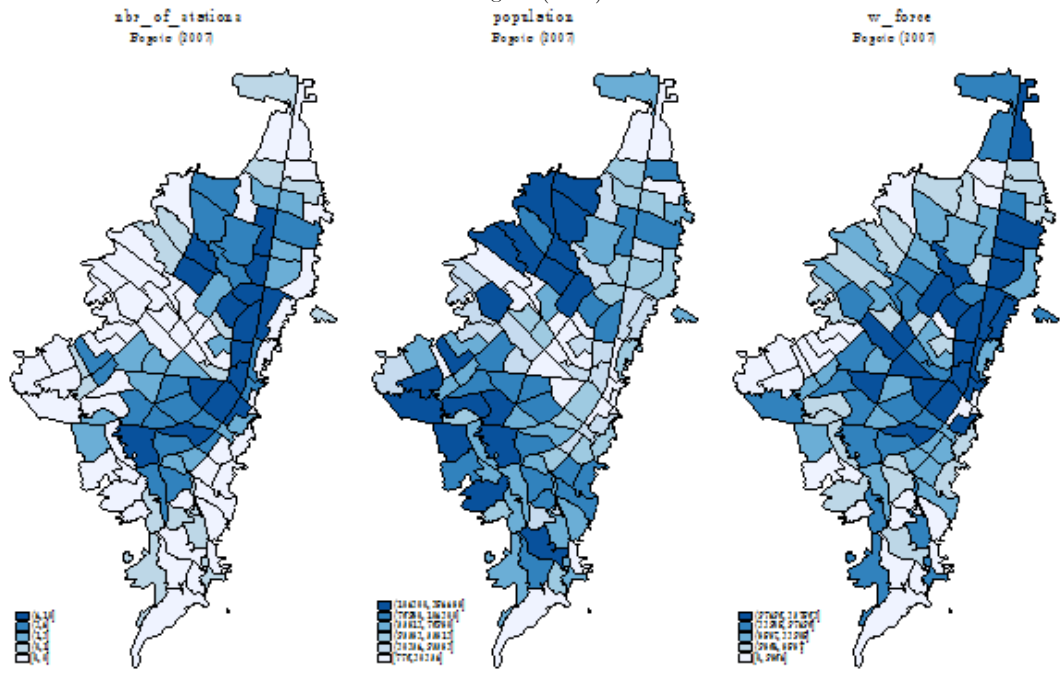
Annex 5: Amenities by Zones (UPZ)

Bogota (2007)



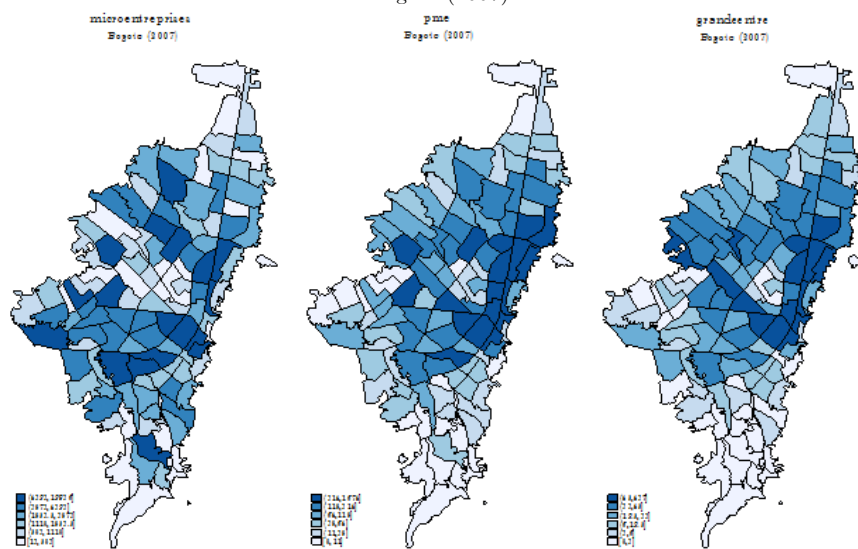
Author calculations

Table 6: Concentration of Stations of TM, Population, Workforce configuration by (UPZ)
Bogotá (2007)



Author calculations

Annex 7: Concentration of enterprises by (UPZ)
Bogota (2007)



Author calculations

7 Bibliography

Attanasi, Giuseppe, Casoria, Fortuna, Centorrino, Samuele and Urso, Giulia, (2012), Cultural investment, local development and instantaneous social capital: A case study of a gathering festival in the South of Italy, No 12-291, TSE Working Papers, Toulouse School of Economics (TSE), <http://EconPapers.repec.org/RePEc:tse:wpaper:25758>.

Bagley Michael N., Mokhtarian, Patricia L. (2002). The impact of neighborhood type on travel behavior: A structural equations modeling approach. *The annals of regional Science*, 36(5), 279-297.

Bocarejo, J. y Oviedo, D. (2010), "Transport Accessibility and Social Exclusion: A Better Way to Evaluate Public Transport Investment?" [en línea], disponible en http://www.ucl.ac.uk/dpu/metrocables/dissemination/Bocarejo_y_Oviedo.pdf, consultado el 27 de septiembre de 2011.

Cao, Xinyu; Mokhtarian, Patricia L.; & Handy, Susan L.(2007). Do changes in neighborhood characteristics lead to changes in travel behavior ? A structural equations modeling approach. *Transportation*, 34(5), 535 - 556. Retrieved from: <http://escholarship.org/uc/item/49393360>

Cervero R, Sandoval O, Landis J (2000) Transportation as a stimulus to welfare-to-work: private versus public mobility. *J Plan Edu Res* 22:50-63

David, Quentin, Janiak, Alexandre and Wasmer, Etienne, (2010), Local social capital and geographical mobility, *Journal of Urban Economics*, 68, issue 2, p. 191-204.

Hansen, W. G. (1959), "How Accessibility Shapes Land Use", en *Journal of American Institute of Planners*, No. 25, pp. 73-76.

Ihlanfeldt, K., and D. Sjoquist. 1991. The role of space in determining the occupation of black and white workers. *Regional Science and Urban Economics* 21:295-99.

Ihlanfeldt, Keith R., Sjoquist, David L. 1998 : "The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform". *Housing Policy Debate* , Vol. 9, Iss. 4.

Kawabata, M. (2003) "Job access and employment among low-skilled autoless workers in U.S. Metropolitan areas." *Environment and Planning A*, Vol. 35, No. 9, pp. 1651-1668.

Krugman, P. (1997), *Development, Geography and Economic Theory*, Cambridge, The MIT Press.

Koenig, G. (1974), “Théorie économique de l’accessibilité urbaine”, en *Revue économique*, Vol. 25, No. 2, pp. 275-279.

Prud’homme, R. y Lee, C. W. (1999), “Size, Sprawl, Speed and the Efficiency of Cities”, en *Urban Studies*, Vol. 36, No.11, pp. 1849-1858.

Shengyi Gao, Patricia L. Mokhtarian, Robert A. Johnston (Exploring connection among job access, empty, income and auto ownership using SEM, 2007)

Wenglenski, S. (2004), “Une mesure des disparités sociales d’accessibilité à l’emploi”, en *Revue l’Economie Régionale et Urbaine (RERU)*, No. 4, pp. 153-169.