Residential Segregation and Public Housing Policy, The Case of Chile

Dionysia Lambiri
Department of Geography, University of Southampton
and
Miguel Vargas
Facultad de Economía y Empresa, Universidad Diego Portales

Abstract

To shed light upon the forces driving Residential Segregation based on income is the aim of the present research. Using a rich micro-database from Santiago of Chile, a transparent methodology has been implemented for understanding the Santiago segregation pattern and the contributions of a set of individuals characteristics in driving it. The analysis has been applied to the years 1992, 1994,1996,1998, 2000, 2003 and 2006, as a way to investigate the Residential Segregation changes over the time. The main results are that an important amount of Residential Segregation is explaining by individuals households characteristics, and that the Chilean housing policy has been an important determinant of Residential Segregation.

1 Introduction

It has been argued that residential segregation based on income has been a side-effect of Chilean Public Housing Policy. However, this hypothesis has not been tested yet. Hence, the aim of this research is to shed light on the forces driving Residential Segregation based on income in Santiago, Chile (**RS** hereinafter). Using a rich micro-database, an easy and transparent methodology, based on Bayer et al. (2004), has been implemented for understanding the Santiago segregation pattern and the contributions of a set of individual characteristics in driving it.

Studying **RS** is relevant because of its potential consequences on the population's well-being. The literature on the subject has indicated that **RS** would have an impact, mainly on joblessness, academic performance, single parenthood, health, the level of criminality, and poverty.[1] Although there are some works that claim that **RS** has no effects (see Cheshire (2007) for a review of this issue), this is still an open questions.

RS sources can be classified as endogenous and exogenous. Endogenous causes are those linked to **RS** that is a consequence of the interaction of individuals' preferences, restrictions and characteristics. **RS** sources can be exogenous when an external force, that does not necessarily have anything to do with individual preferences, sorts people. Regarding this kind of sources, a taxonomy of two groups of drivers has been proposed: Those related to policies which either intentionally or unintentionally produce segregation, and real estate markets dynamics. The latter is understood to be exogenous from the standpoint of individuals who are consumers of housing services. With respect to those drivers related to public policy, it is important to have a clear idea of their contribution to the generation of **RS** because if they are significant, the policymakers can redesign, or even change, the public policy in order to attenuate **RS**, something that is particularly important when segregation is based either on race or income, which is patently evident, given their negative effects on the well-being of individuals, the most pernicious kind of **RS**. A good example of an intentional public policy, or *de jure* segregation, is South African apartheid (for a deeper analysis of the effects of apartheid on racial segregation in South Africa see Christopher (1990)). But there is no need to deal with such a drastic legislative programme to find these sort of external forces working to generate **RS**, as

[1] See for instance Dawkins et al. (2005); Charles et al. (2004); Clapp and Ross (2004); LaVeist (2003); Dosh (2003); Burton (2003); Yinger (2001); Massey (2001); Madden (2001); Wilson and Hammer (2001); Logan and Messner (1987); Burnell (1988) and King and Mieszkowski (1973).

in the case of some public policies that unintentionally generate **RS**. For instance, there is evidence that the Chilean Public Housing Policy is one of the causes of **RS** in this country. In order to reduce the accumulated housing shortage and to control the slums and poverty-stricken urban settlements, the Chilean government implemented large-scale housing production during the 90s.[2] This policy was relatively successful in achieving its main goal of reducing the housing deficit from 771,935 to 543,542 (Castillo, 2008; Ruprah and Marcano, 2008; Sugranyes, 2004) but it generated **RS** as a collateral effect. The reason is that in order to keep housing units affordable, the new neighbourhoods were built on cheap land located far from the city centre, mainly on the southern boundary of the city. Tokman (2006) shows that households with a similar socio-economic background that have had freedom of choice to select their housing units are located, on average, one kilometre closer to the city centre than those households living in the new affordable neighbourhoods. Consequently, this policy —despite its success in reducing the housing deficit—seems to be an important driver of **RS** in Santiago. The latter is a hypothesis that has not been tested yet, although this research aims to do so.

A fully qualitative and quantitative comprehension of the causes of **RS** is still a pending task. Firstly, there is no theoretical consensus on the determinants of **RS**. Although the interactions among individual based models, such as Schelling's, provide a rich research framework, there is a lack of mathematical formalization of these models, as Krugman (1996); Young (1998) and Zhang (2004) have pointed out; the econometric difficulties that must be dealt with, mainly, identification of problems and the scarcity of micro-databases at lower aggregation levels have made empirical analysis a difficult task.

Consequently, the case would be particularly interesting to research, due to the fact that there is a good micro-database at a census tract level, allowing for the exploration of the impact of individual characteristics on **RS**. The latter helps to shed a new light on the forces driving **RS**. Santiago is also interesting because, according to Rodríguez (2001), there is evidence indicating that it is one of the most segregated cities in the world based on income terms.

[2] For a history of Chilean Housing Policy see Hidalgo (2002)

Methodologically, this research is based on Schelling (1971) which identifies a group of different mechanisms indirectly related solely to race that can be the cause of racial segregation. Something similar could happen with **RS** based on income, where different variables linked to income, such as age, gender, marital status, housing prices, or whether housing is affordable, can explain the level of observed **RS**. Therefore, in order to understand **RS** it is important to understand how the characteristics of households can drive it.

An important novelty of this research lies in the fact that this is the first time that a procedure like this has been used to study economic **RS**. The use of an alternative estimation procedure is also proposed in order to avoid the difficulties arising from fractional response variables, like the case studied here where the dependent variable will be a index number between 0 and 1.

The article is organized as follows: the next section describes Santiago's main characteristics. Section 2 explains the housing policy in Chile. Then Section 3 explains the data set used in the analysis. Section 4 describes some **RS** measurements and Santiagos **RS** pattern. Section 5 explains the methodology used and then explores the extent to which the correlation of households' characteristics can explain the observed pattern of economic **RS** in Santiago. Section 6 concludes.

1.1 The City of Santiago

Santiago is Chiles capital and principal city. It is located in the country's central valley, at an elevation of 520 m. It comprises a conurbation of 32 municipalities that altogether are called "Greater Santiago". Each one of these municipalities is independent, with its own mayor —elected by voters every four years—. Each has its own public medical services, schools and regulations and collects its own municipal taxes. Besides, there is a central government authority: the Governor of Santiago. The municipality of Santiago Centro, located in the centre of the Greater Santiago, encompasses the executive and judicial branches.

Figure 1 shows Santiago's location and figure 2 shows a map of the city with the location of the municipalities.

Figure 1: Santiago location

1.2 Geography

The city lies in the centre of the Santiago Basin, a large bowl-shaped valley consisting of a broad plain surrounded by mountains. It is flanked by the main chain of the Andes on the east and the Chilean Coastal Range on the west. Its northern boundary is the Cordn de Chacabuco, a transverse mountain range of the Andes, whereas its southern boundary is Angostura de Paine, where an elongated spur of the Andes almost reaches the Coastal Range. The Santiago Basin is part of the Intermediate Depression and is remarkably flat, interrupted only by a few hills.

Figure 2: Santiago municipalities

1.3 Climate

Santiago has a mild Mediterranean climate: relatively hot dry summers (November to March) with temperatures of up to 35 degrees Celsius on the hottest days; winters (June to August) are more humid with cold mornings, typical maximum daily temperatures of 15 degrees Celsius, and minimum temperatures of a few degrees above zero. Occasional snowfalls occur and may extend throughout the city. Mean rainfall is 360 mm per year and is heavily concentrated in the cooler months.

1.4 Pollution

Thermal inversion causes high levels of smog and air pollution to be trapped and concentrated within the Central Valley during the winter months. In the 1990s air pollution fell by about one-third, but there has been little progress since the year 2000.

As of March 2007, only 61% of the waste-water in Santiago was treated; this increased to 71% by the end of the same year. However, the Mapocho river, which crosses the city from the north-east to the south-west of the Central Valley, remains contaminated by household, agricultural and industrial sewage, and by upstream copper-mining waste which is dumped unfiltered into the river (there are a number of copper mines in the Andes range east of Santiago).

Noise levels on the main streets are high, mostly because of noisy diesel buses. Diesel trucks and buses are also major contributors to winter smog. A lengthy process of renovating the bus system began in 2005 and will last until 2010.

1.5 Demographics

Santiago has about 5.9 million inhabitants and the average rate of demographic growth per year is 1.5%. The city has been expanding at the rate of 3.4% per year, on average, over the last 40 years. A migration process from the central areas of the city to peripheral areas is one of the explanations of this phenomenon.[3] Mattos (2002) shows that between the years 1992 and 2002 the central areas have lost, on average, 7.4% of their population whereas the population of peripheral areas has increased by 47.4% on average. Particularly remarkable are the Maipú and Quilicura growth rates, as can be seen in table 1.

1.6 Urbanization

Santiago is a highly urbanized city - about 70% of the citys surface area. In general the municipalities located in the city centre, such as Santiago Centro, Quinta Normal and La Cisterna, are fully urbanized, whereas most of the still unurbanized land is located on the

[3] See INE (2005).

Table 1: Population

Municipality	Population 1992	Population 2002	Annual Variation (%)
Cerrillos	72,649	71,906	4.41
Cerro Navia	155,735	148,312	-0.10
Conchalí	152,919	133,256	-1,37
El Bosque	172,854	175,594	0,16
Estación Central	140,896	130,394	-0,77
Huechuraba	61,784	74,070	1.83
Independencia	77,794	65,479	-1.71
La Cisterna	94,712	85,118	-1.06
La Florida	328,881	365,674	1.07
La Granja	133,285	132,520	-0.06
La Pintana	169,640	190,085	1.14
La Reina	92,410	96,762	0.46
Las Condes	208,063	249,893	1.85
Lo Barnechea	50,062	74,749	4.09
Lo Espejo	120,075	112,800	-0.62
Lo Prado	110,933	104,316	-0.61
Macul	120,708	112,535	-0,70
Maipú	256,550	468,390	6.20
Ñuñoa	172,575	163,511	-0.54
Pedro Aguirre Cerda (PAC)	130,441	114,560	-1.29
Peñalolén	179,781	216,060	1.86
Providencia	111,182	120,874	0.84
Pudahuel	137,940	195,653	3.56
Quilicura	41,121	126,518	11.89
Quinta Normal	116,349	104,012	-1.11
Recoleta	164,767	148,220	-1.05
Renca	128,972	133,518	0.35
San Joaquín	114,017	97,625	-1.54
San Miguel	82,869	78,872	-0.49
San Ramón	100,817	94,906	-0.60
Santiago Centro	230,977	200,792	-1.39
Vitacura	79,375	81,499	0.26

Source: National Institute of Statistics periphery (see table 2). The composition of urbanized land varies considerably between one municipality and another. Some of them are mainly residential —such as Las Condes, Lo Barnechea and Providencia—, and some of them are mainly dedicated to commercial and industrial activities —such as Quilicura, Cerrillos and Pudahuel—. Altogether, more than the 50% of the citys surface area is residential (Galetovic and Poduje, 2006).

1.7 Density

In general terms, density is the number of people living on a given piece of land. As Galetovic and Poduje (2006) point out, there are several ways to define this concept. The most general is population divided by municipal surface area (column 1 in table 3). However, this way of measuring density can generate some difficulties because a considerable part of the surface area of some municipalities falls outside the urban boundary(for example Lo Barnechea and Maipú). Consequently, urban density is commonly described as: the area within the urban boundary divided by the total population (column 2 in table 3). Using this calculation, one can see that urban density is ten times normal density. Urban density has some drawbacks as well. The main one is that it can underestimate the real degree of concentration, due to the fact that in some countries, for instance, a significant part of the surface area is unurbanized. Furthermore, urban density increases when urban boundaries expand. A measure devoid of these problems is the adjusted density: the total population divided by the urbanized area (column 3 in table 3). This density is even higher than the previous ones. Finally, if ones purpose is to measure the degree of concentration of people more accurately, the best index to use is residential density: the total population divided by the residential area (column 4 in table 3). The residential density is twice the urban density and varies more between municipalities. The case of Providencia is interesting. Its adjusted density is quite similar to the average adjusted density, but its residential density is very high. This is because a significant part of Providencias surface area is dedicated to commercial activities. Something similar occurs in Recoleta.

Table 2: Area and distance to town

Municipality	Area (ha)	urbanized area (ha)	Distance to town (km)
Cerrillos	1,696	1,350	11.8
Cerro Navia	1,118	907	8.6
Conchalí	1,103	1,103	7.5
El Bosque	1,429	1,428	14.9
Estación Central	1,433	1,388	6.4
Huechuraba	4,461	1,156	7.8
Independencia	745	745	3.2
La Cisterna	1,000	1,000	10.9
La Florida	7,113	3,785	16.0
La Granja	1,008	1,008	14.3
La Pintana	3,050	1,661	16.5
La Reina	2,369	1,769	12.1
Las Condes	9,991	3,818	11.6
Lo Barnechea	101,258	2,577	19.3
Lo Espejo	822	822	12.7
Lo Prado	659	659	6.5
Macul	1,284	1,284	10.4
Maipú	13,713	4,461	14.9
Ñuñoa	1,690	1,690	7.7
Pedro Aguirre Cerda (PAC)	868	868	10.1
Peñalolén	5,338	2,377	13.3
Providencia	1,439	1,291	5.2
Pudahuel	19,679	2,882	10.2
Quilicura	5,725	2,210	12.7
Quinta Normal	1,190	1,190	5.2
Recoleta	1,584	1,355	4.4
Renca	2,332	1,393	8.8
San Joaquín	1,001	1,001	8.3
San Miguel	964	964	7.1
San Ramón	631	631	13.9
Santiago Centro	2,311	2,311	-
Vitacura	2,770	1,972	12.1

Source: Galetovic and Poduje (2006)

Table 3: Density

Municipality	Density	Urban density	Adjusted density	Residential density
Cerrillos	42	42	53	164
Cerro Navia	133	133	164	215
Conchalí	121	121	121	190
El Bosque	123	123	123	189
Estación Central	91	91	94	170
Huechuraba	17	36	64	156
Independencia	88	88	88	143
La Cisterna	85	85	85	116
La Florida	51	85	97	140
La Granja	132	132	132	204
La Pintana	62	82	114	148
La Reina	41	50	55	83
Las Condes	25	57	65	106
Lo Barnechea	1	18	29	47
Lo Espejo	137	137	137	218
Lo Prado	158	158	158	224
Macul	88	88	88	154
Maipú	34	85	105	185
Ñuñoa	97	97	97	145
Pedro Aguirre Cerda (PAC)	132	132	132	193
Peñalolén	40	66	91	135
Providencia	84	84	94	214
Pudahuel	10	87	68	195
Quilicura	22	35	57	183
Quinta Normal	87	87	87	119
Recoleta	94	94	109	216
Renca	57	57	96	179
San Joaquín	98	98	98	173
San Miguel	82	82	82	151
San Ramón	150	150	150	213
Santiago Centro	87	87	87	172
Vitacura	29	35	41	72

Source: Galetovic and Poduje (2006)

2 Chilean Public Housing Policy

The main purpose of Chilean Public Housing policies has been to reduce the lack of affordable dwellings and to increase ownership. Accordingly, since the second half of the 20th century the Ministry of Housing has focused its efforts on the mass construction of affordable dwellings. As a result, the Ministry of Housing has built three quarters of all housing units built in Chile, as can be appreciated in Figure 3.

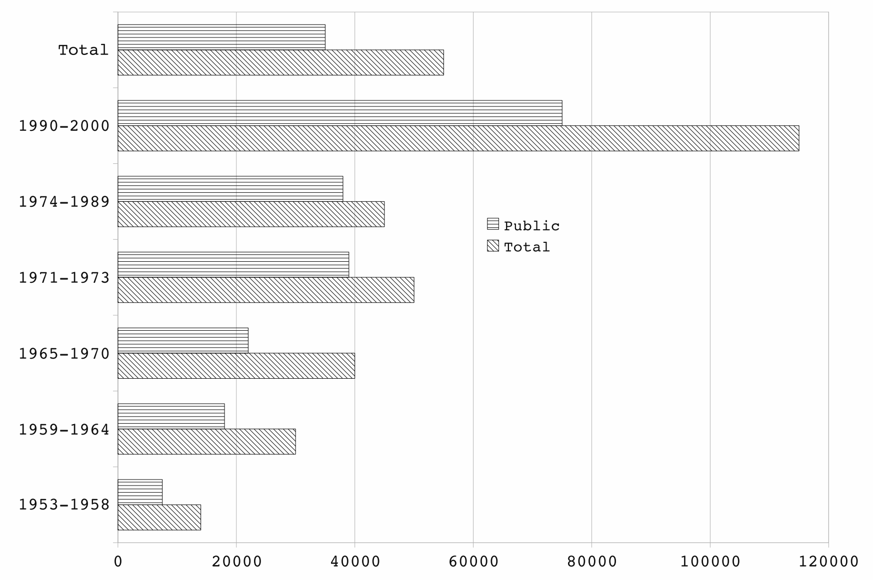

Figure 3: Average number of housing units built per year (Tokman, 2006)

If this policy is assessed as a means for reducing poverty, because it has been used as an income distribution programme, there is no doubt that it has been successful. The latter was particularly true throughout the nineties, a period in which this policy was very aggressively applied. As a matter of fact, more than 400,000 housing units were built during the 90s, either by the State or by means of subsidies provided by the State.

Consequently, during the last few decades, an important increase in the stock of housing units has been observed, enabling most households to access housing solutions. In fact, according to the 2002 census, 76%of households owned the housing units they lived in, a rate much higher that the average of the OECD countries which is 68% (OECD, 2004). However, as Simian (2010) points out, Latin American countries have always exhibited higher rates of ownership, 72% on average. A relevant fact related to the ownership rate is that it has been evenly distributed among income quintiles. Using the 2006 CASEN survey Simian (2010) calculates ownerships rates, which are 68%, 70%, 69%, 70% and 68% for the first, second, third, fourth and fifth quintile, respectively. Another important achievement of this policy has been the improvement in the quality of housing units, in terms of overcrowding. According to the Ministry of Housing, there were 242,000 secondary households living in main household dwellings in 2002, whereas this figure dropped to 126,000 in 2006.

This policy rests on two main pillars: subsidies on demand, in the form of a voucher, and subsidies for the construction of affordable dwellings. The first group of subsidies focuses on households with savings capacity and access to loans . There are other kinds of subsidies, namely: the General Unified Subsidy (SGU), the Special Programme for Workers (PET), the Rural Subsidy, the Freedom of Choice modality for Basic and Progressive housing and the D.S. 40 Subsidy. The second group focuses on the construction of affordable housing for low-income families without savings capacity. These subsidies include the Basic and Progressive housing programmes. Table 4 shows the relative importance of each subsidy programme, and the next subsections explain these two subsidies schemes in greater detail.

The hypothesis that is tested here is that the affordable housing policy has been an important driver of residential segregation, so the main focus is on trying to understand how this policy works; however, in order to have a better idea of the entire system, a full description of it will be given in the following subsections, which mainly follow Simian (2010).

Table 4: Housing built by subsidy programme

Programmes	1990-1999	2000-2007
Affordable Housing:		
Basic Housing	27.7%	10.1%
Progressive Housing	6.0%	0.7%
Dynamic Social Housing without debt		3.2%
Vouchers and other subsidies:		
SGU	27.9%	6.8%
PET	18.9%	9.8%
Rural	8.6%	8.4%
Private Basic	3.0%	5.7%
Progressive I	5.7%	6.0%
Progressive II	1.7%	1.0%
Leasing	0.5%	2.3%
Solidarity Fund for Housing I		18.0%
Solidarity Fund for Housing II		1.9%
Location Subsidy		0.0%
D.S. 40		13.0%
Households heritage protection		4.7%
Housing Maintenance		9.2
	100%	100%

Source Simian (2010)

2.1 Affordable housing

The subsidies for affordable housing mainly focus on households with a total income below US$ 400 per month. Housing units have a surface area of between $14m^2$ and $40m^2$, depending on the households' income. The housing units are small because they have been designed as either progressive or basic units which can be extended and improved by households once they have received them. To qualify as a beneficiary, a household must have prior minimum savings, which vary depending on the kind of subsidy it will apply for, which in turn depends on its income. The Basic Housing programme was created in 1984. The average size of this kind of housing unit is $40m^2$. The units were initially houses, but after a while the programme switched to flats as a means of reducing land costs. The average price of these units is US$10,000. The subsidy in this case is US$6,000. The progressive housing scheme was created in 1990. This kind of units consist of one toilet, one kitchen and one room. They have an average size of $14m^2$ on a $100m^2$ plot. The price of these units is US$ 6,000, and the subsidy in this case is US$ 5,000.

The cost not covered by the subsidy must be paid by the households using their accumulated savings. For instance, the poorest households earning less than 114 US$ per month, will receive a $14m^2$ unit whose price is US$ 6,000. The household must have accumulated savings of US$ 300, and thereafter the State will provide a subsidy of US$ 5,000. Households that are capable of paying a larger deposit and applying for a mortgage, and households with an income of more than 227 US$ per month, can access houses of $40m^2$ whose price is US$ 10,000. In this situation households must have prior savings of US$ 1,000 and finance US$ 3,000 throughout a mortgage. The remaining US$ 6,000 corresponds to the subsidy that this kind of households will receive. The average number of basic and progressive housing units that were built between 1990 and 2006 was 21,000 per year. 96% of them were basic housing units, which means that progressive housing were never important in the affordable housing policy.

According to Simian (2010), basic housing units are concentrated in the two poorest quintiles, which have received more than 50% of all subsidies, which means that this policy has been successful in focalizing its efforts on the poorest segment of society.

A critical issue that arose was the increasing number of delinquent mortgages. Besides, due to the lack of private mortgage offers, the Ministry of Housing provided these loans itself. The latter had a momentous impact on housing policy funding. As a result, since 2000 the Ministry of Housing has focused only on housing units that are fully funded by the households savings and the ministrys subsidy.

The affordable housing policy has been very successful in providing solutions to poor households. Nevertheless, its application has brought some serious problems with it. Among them, the need to buy cheap land on urban boundaries for building the solutions, has created problems based on income segregation. Despite its relevance this problem has never addressed in a quantitative way, for gauging how important it has been as a segregation driver.

Figure 4: Subsidies paid (source: Ministry of Housing)

2.2 Vouchers and other subsidies

The main aim of these subsidies is to reduce the asymmetric information problem between low and middle income families, and financial institutions. The voucher is awarded for buying a property. It was introduced in 1978 and after several changes, it became the Unified General Subsidy (SGU). It was discontinued in 2004. There are also two other subsidies: rural subsidies, and the Special Programme for Workers (PET), which is for groups of households instead of just one household. However, the most important one by far has been the SGU. As a fact of matter, according to the Ministry of Housing data, 90% of the subsidies awarded between 1990 and 2005 were SGU.

These subsidies were awarded solely for acquiring new housing units until 1995, but since 1996 they could be used for buying a second housing unit as well.

These subsidies were not focus on the poorest households, because they have a very low saving capacity. Hence, one would expect that their main recipients would be households in the third quintile. However, using the 2006 CASEN survey, Simian (2010) found that 50% of them were awarded to households in the two richest quintiles. One reason is that the applicable housing price cap was relative high, about US$65,000, so that middle income households could apply for these subsidies.

As previously mentioned, these subsidies were phased out in 2004 and were replaced by the D.S. 40 subsidy, which, together with other policy changes, is explained in the next subsection.

2.3 The new courses of the housing policy

Since 2000 a series of important changes were introduced to the Chilean housing policy, the main ones being the following: first of all, the amount of money given as a subsidy and the applicable housing price cap have been reduced, as a way of improving means testing; secondly, the Ministry of Housing no longer builds affordable housing units. This has now been entrusted to private building companies, through a bidding process. Due to the above, private institutions must gauge the demand for developing an affordable housing project. Thirdly, the Ministry of Housing no longer gives any loans.

New subsidy programmes have also been implemented. The main ones are subsidy D.S. 40 and the Solidarity Fund for Housing. Subsidy D.S. 40 replaces the SGU subsidy. It was created in 2004. The main difference with its predecessor is that the maximum price of the housing unit has been reduced, so households, depending on their income, can apply for units whose price is lower than US$ 27,000 or lower than US$44,000. The amount of the subsidy awarded cannot be greater than US$4,000, which implies an important subsidy reduction. Since 2004, more than 50,000 of this kind of subsidies have been awarded.

The Solidarity Fund for Housing was designed to contribute to the construction of new housing units or the purchase of second hand units, both of them with a price, on average lower than US$ 21,000 (the price can vary depending on the municipality in which the units are located). In order to apply, households must have prior minimum saving of US$440. This subsidy focuses on the poorest segment of households. According to Simian (2010) in 2006 more than the 70% of these subsidies were given to the two poorest quintiles.

2.4 Affordable housing policy and segregation

The argument stating that the Public Housing Policy has generated **RS**, is based on the fact that the Ministry of Housing, when choosing the location of an affordable housing unit project, as Tokman (2006) points out, takes only the following criteria into account.

First of all, it only considers the land prices and building costs, disregarding significant costs such as investments and costs in which the local authorities will have to incur, such as garbage collection, public lighting, green area maintenance, etc. Besides, there are others costs related to the new infrastructure that must be built, such as schools, hospitals and police stations that must be covered by other branches of government, and finally, there are costs such as the externalities that the new projects generate, and commuting costs that must be paid by households.

Secondly, the Ministry of Housing gives priority to big projects. Hence, the projects require large expanses of land.

Thirdly, the prices of housing units cannot be greater than the maximum allowed. All the tenders consider a maximum price per unit and the land price cannot be more than 21% of the total cost. The latter requirement restricts land selection to the cheapest areas, which are mainly located on the city boundaries and almost all of them are unurbanized.

Fourth, before 2002, the Ministry of Housing did not consider the project environment and location. The main consequence of these criteria used by the Ministry of Housing is that almost all affordable housing units have been built on the urban periphery, because land is cheaper in these areas and extensive expanses of land are vailable. As a matter of fact, Tokman (2006) shows that in 2002 the biggest concentration of projects within Greater Santiago were located in Quilicura, La Pintana, Renca, Pudahuel and Peñalolén. Hence, isolated neighborhoods inhabited by low income households are located on the southern and northern boundaries of the city. Evidence indicates that the Public Housing Policy has been the most important **RS** driver; however, this hypothesis is yet to be tested, and is the main purpose of this research. Besides, it is relevant to cast a light on the extent that this policy can affect **RS** in Santiago. Both issues mentioned above can have important policy implications.

3 Data

The purpose of the Ministry of Planning and Cooperation (**MPC**) is to design and apply Chilean development policies and programmes at a national and regional level. Besides, it must set out the goals of public investment and assess them, among other responsibilities. In order to achieve these tasks **MPC** has a variety of tools; one of them is the National Socioeconomic Characterization Survey (CASEN for its Spanish acronym). Based on the World Bank living standards measurement survey methodology, CASEN contains nearly 300 variables describing individuals and families within the basic domains of household characteristics, education, health, employment and income. Conducted in November, the CASEN surveys provide the basic information needed to help government planners reduce poverty, improve the standard of living, and help integrate excluded sectors into the development process.

This survey has been carried out every two years since 1985, except in 1989 and 2002, when it was forwarded to 1990 and 2003, respectively). The **MPC** is responsible for this survey.

CASENs main purpose is:

1. to measure the redistributive impact of the Government's social expenditure;
2. to measure the income distribution corrected by Government transfers;
3. to evaluate social programmes;
4. to describe the population according to income, level of education, housing conditions and jobs;
5. to measure the level of poverty and to link poor households with the other variables the survey measurements.

3.1 The Sample

The survey sampling is geographically stratified. There are urban and rural strata. In the 2003 survey, 553 strata were used. The sample size was 68,400 households with information on 272,000 individuals, representing about 1.8% of the total population and 1.7% of all households. The sampling error is 0.4%. The latter information is valid for the whole country. For the specific case of the present analysis, Santiago's samples for 1992, 1994, 1996, 1998, 2000, 2003 and 2006 were considered. All of them contain information for 32 municipalities, the total number of Santiago's municipalities, on average, 8,861 households and 620 census tracts. In all these cases, only the urban strata information was considered.

3.2 The questionnaire

The questionnaire is very similar to the census questionnaire, but it is longer and considers income, unlike the census. The information is classified in the 6 following modules:

1. Residents: in this module each household member is described on the basis of kinship, sex, age, disabilities, equity, access to technologies, membership of government social programmes, and ethnicity.

2. Education: this module contains information on literacy, such as: do household members attend any kind of school? If not why? Do they plan to do so in the next year (to study, to work and so on), level of education, kind of education (public, private or subsidized) and if any of the household's member has a scholarship.

3. Health: the health module contains information about the kind of health care (public or private), conditions and medicines, and if any of the household members receive food or medicines from the Government.

4. Employment and income: if any of the household members are 12 or older, they have to answer whether they worked the last week or whether they have been looking for a job within the last two months; if they are looking for a job is it the first time? If they are not looking for a job, why? If they have a job, is it their first job? What kind of job is it? How long have they been working? How much do they earn? What kind of pension system do they have? What were the household members doing in the last survey year?

5. Other sources of income: this module concerns income from rents, savings, transfers, retirement and subsidies.

6. Housing: this module concerns the tenure characteristics, the price and quality of the housing unit and mortgage information.

4 Santiago Pattern of Residential Segregation

In this section the level **RS** in Santiago is calculated, but previously a brief explanation about some measurements of **RS** is given.

4.1 Measurements of Segregation

Duncan and Duncan (1955) is the first systematic analysis of segregation indices. According to them there is not information in any of the segregation indices that the index of dissimilarity does not contain. James and Taeuber (1985) developed a set of criteria to evaluate segregation measurements: organizational equivalence, size invariance, transfers and composition invariance. They also demonstrate that indices that are highly correlated in empirical studies may, nonetheless, behave different under certain circumstances, such as when the population shares of groups change. Another important research on **RS** indices is Massey and Denton (1988). In this work they classify segregation indices, using factor analysis, into five dimensions: evenness, exposure, concentration, centralization and clustering. For each of these dimensions they determine which is the best index.[4] These indices can also be classified depending on their spatial feature. So there are non-spatial indices and spatial indices. Evenness, exposure, concentration and centralization indices belong to the first group and the clustering index belongs to the second one. There are some other indices[5]; however, thus far, these have been the most important in literature. Another interesting way to measure exposure is proposed by Bayer et al. (2004). The strength of this index relies on the fact that is very easy to calculate and to interpret. Also, it can be divided into several meaningful ways. Now a brief description of all indices that have been used in the present research for measuring Santigo's **RS**.

[4]One problem with this research is that does not consider measurements of multi-group segregation. Instead, all the analysis is conducted just considering dichotomous measurements. Reardon and Firebaugh (2002) apply the criteria of James and Taeuber (1985) to the case of multi-group measurements. They derive and evaluate a set of six multi-group segregation indices. They conclude that the information theory index is the most, conceptually and mathematically, satisfactory index .

[5]See Reardon and Firebaugh (2002).

1. **Evenness** refers to the differential distribution of two social groups among geographical units in a city. A minority group is said to be segregated if it is unevenly distributed over units. Evenness is maximized and segregation minimized when all units have the same relative number of minority and majority members as the city.

   The most appropriate index to measure evenness is the dissimilarity index:

$$D = \sum_{i=1}^{n} \left[\frac{t_i}{p_i} - \frac{P}{2TP(1-P)} \right] \tag{1}$$

   where t_i and p_i are the total population and minority population of unit i, and T and P are the population size and minority proportion of the whole city.

2. **Exposure** refers to the degree of potential contact, or the possibility of interaction, between minority and majority groups within geographic area or city.

   The most appropriate index for measuring exposure is the exposure index:

$$P = \sum_{i=1}^{n} \left[\frac{x_i}{X} \frac{y_i}{t_i} \right] \tag{2}$$

   As mentioned before, an alternative way for measuring exposure, proposed by Bayer et al. (2004), is:

$$E(r_j, R_k) = \frac{\sum_i r_j^i R_k^i}{\sum_i r_j^i} \tag{3}$$

   where R_k^j is the fraction of households belonging to the group k in household i's neighbourhood, and r_j^i is a dummy variable taking value one if household i belongs to group k. This index can be interpreted as the average exposure of households belonging to group i to households belonging to group k.

4.2 Santiago's RS

The 32 municipalities of Santiago are different in aspects such as geography, demographics, economics, public infrastructure and average income. The east end of the city (historically called "the high quarter") corresponds to the highest average income area. In the city centre there is a group of middle class municipalities, meanwhile in the southern and north-western ends the poorest municipalities are located.

These differences can be very important. For instance, although infant mortality in Chile is low (7.9 deaths/1,000 live births), there is a difference of 75% between the municipality with the lowest rate of mortality (excluding La Cisterna, a sort of out-layer, which is a low income municipality with a low infant mortality rate), Lo Barnechea, one of the richest municipalities in Santiago, and the municipality with the highest one, Lo Espejo, one of Santiago's poorest municipalities (see table 5). The level of illiteracy is also a good example (table 6). The difference between Vitacura and La Pintana in this index is 95%. There are some municipalities with almost no poverty, such as Vitacura, meanwhile there are others where the level of poverty is 30%, like La Pintana (see table 7). These differences, among others factors, have an important impact on the funds available for municipalities, because a significant part of them come from municipality taxes. Consequently, richest municipalities have better public facilities. All these facts have an important impact on land price, something that literature has identified as one of the determinants of **RS**. This phenomenon is a poverty traps driver as well.

The latter suggests that Santiago would be a segregated city. Rodríguez (2001), using the census of 1982 and 1992, concludes that Santiago is a segregated city, but this segregation is moderated and it has been reduced. The problems with this conclusion are mainly two. First, the idea that the level of segregation is not too high come from comparing the income segregation of Santiago with international evidence of racial segregation. This is a mistake because the levels of racial segregation are systematically higher than the levels of income segregation. Abramson et al. (1995) shows that the average Dissimilarity Index value for the 100 largest cities in US in 1990 based on income is 0.36, meanwhile the average value of the same index based on ethnicity is 0.61. Therefore if one compares Rodríguez (2001) results, a Dissimilarity Index value of 0.47 in 1982 and 0.4 in 1992, with international evidence for income segregation, the conclusion is that Santiago is highly segregated in economic terms. The second drawback arises from the conclusion that seg-

Table 5: Infant mortality rate (deaths/1,000 live births)

Municipality	%	Municipality	%
La Pintana	12.2	La Florida	12.9
El Bosque	14.6	Maipú	16.2
La Granja	13.7	Estación Central	16.7
Cerro Navia	11.8	Santiago Centro	14.4
Renca	14.3	Macul	10.3
Huechuraba	13.8	Cerrillos	8.1
Recoleta	13.9	Quinta Normal	14.0
San Ramón	14.8	Independencia	10.2
Lo Espejo	19.7	La Cisterna	4.0
Quilicura	12.4	San Miguel	14.7
Lo Prado	18.4	Lo Barnechea	5.6
Peñalolén	12.9	Las Condes	11.3
San Joaquín	11.8	Ñuñoa	13.6
Conchalí	12.8	La Reina	14.1
Pedro Aguirre Cerda	14.7	Providencia	7.5
Pudahuel	12.6	Vitacura	6.7

Source: National Institute of Statistics

Table 6: Rate of illiteracy (%)

Municipality	%	Municipality	%
La Pintana	5.5	La Florida	2.0
El Bosque	2.4	Maipú	0.8
La Granja	3.6	Estación Central	2.0
Cerro Navia	3.2	Santiago Centro	0.5
Renca	3.5	Macul	2.9
Huechuraba	2.2	Cerrillos	1.4
Recoleta	1.4	Quinta Normal	1.3
San Ramón	2.9	Independencia	0.6
Lo Espejo	3.2	La Cisterna	1.6
Quilicura	0.9	San Miguel	0.4
Lo Prado	2.6	Lo Barnechea	1.9
Peñalolén	4.4	Las Condes	0.6
San Joaquín	1.5	Ñuñoa	0.4
Conchalí	1.8	La Reina	1.7
Pedro Aguirre Cerda	2.4	Providencia	0.4
Pudahuel	2.2	Vitacura	0.3

Source: CASEN 2006

Table 7: Poverty in Santiago by Municipality

Municipality	%	Municipality	%
La Pintana	28.86	La Florida	14.22
El Bosque	26.91	Maipú	13.42
La Granja	24.91	Estación Central	13.1
Cerro Navia	24.04	Santiago Centro	12.31
Renca	22.64	Macul	12.18
Huechuraba	21.69	Cerrillos	11.59
Recoleta	20.81	Quinta Normal	11.35
San Ramón	20.11	Independencia	11.09
Lo Espejo	19.94	La Cisterna	10.25
Quilicura	18.55	San Miguel	7.93
Lo Prado	18.07	Lo Barnechea	5.41
Peñalolén	17.91	Las Condes	4.34
San Joaquín	17.51	Ñuñoa	3.48
Conchalí	16.05	La Reina	2.26
Pedro Aguirre Cerda	15.63	Providencia	0.5
Pudahuel	15.19	Vitacura	0.27

Source: CASEN 2006

regation has been diminished by only comparing two points in time. This is clearly not statistically significant.

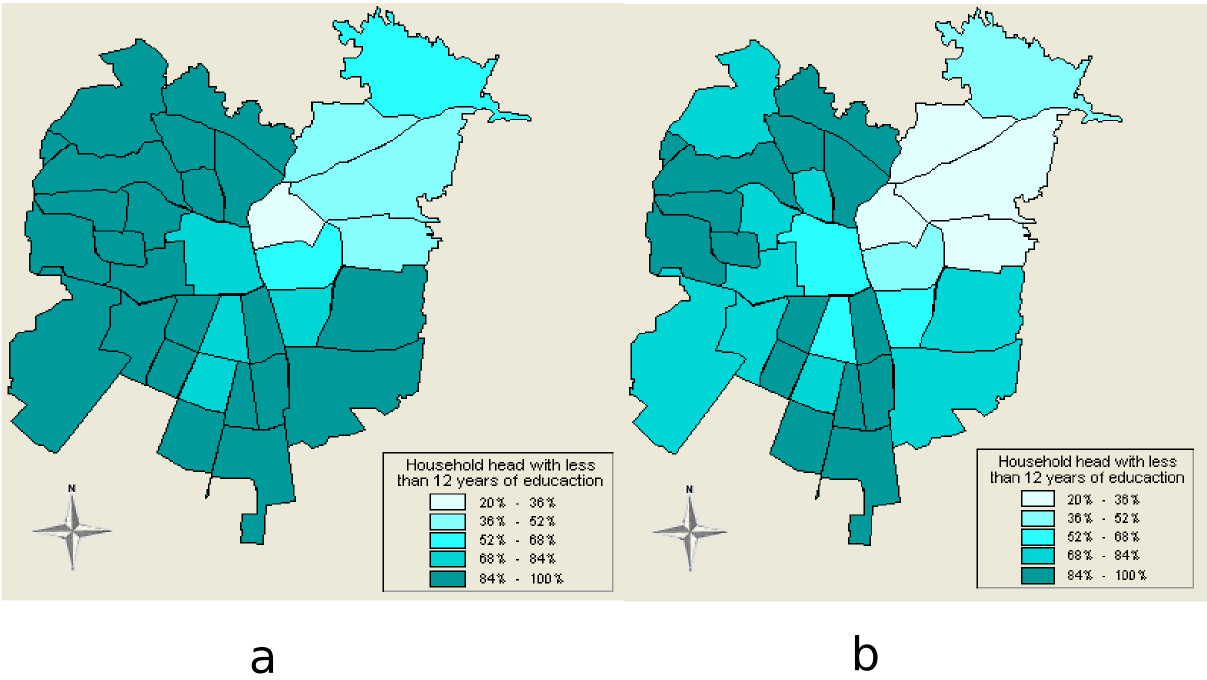

Figure 5: Household heads level of education (CASEN 1992 and CASEN 2003)

Using a different database, the dissimilarity, isolation and exposure indices were calculated for years 1992, 1994, 1996, 1998, 2000, 2003 and 2006 (see the section on data description). The household head educational achievement was used for classifying families. In particular, families have been split into two groups: household head with less than 12 years of education and household head with more than 12 years. The reason is that in the Chilean educational system the elementary and high school phases together sum up to 12 years; therefore, having more than 12 years of education implies some sort of professional education. The level of education is used because is more correlated with permanent income than direct information about income, which is closer to current income. Figure 5 shows differences of this variable among municipalities in 1992 (Figure 5.a) and 2003 (Figure 5.b). The north-western end of the city (Providencia, Ñuñoa, Las Condes, Vitacura, La Reina and Lo Barnechea) concentrates families with the highest levels of educational achievements, i.e. the better-off families. This phenomenon does not change very much in 2003 respect to what happens in 1992. However, the municipality percentage of families head with a low level of education is lower in 2003. The reason is a general improvement of families educational achievement, as a result of a successful policy for reducing poverty (people living under poverty line in 1990 were 40%, meanwhile in 2003 this percentage was 17%). Nonetheless, the latter does not imply a segregation reduction. Figure 6 shows the level of exposure (the opposite to isolation) in 1992 (Figure 6.a) and 2003 (Figure 6.b) in each municipality. As it can be seen, exposure is low (less than 0.025) in 53% of the municipalities in 1992, and in 56% of the municipalities in 2003.

Table 8: Households head with low education

Municipality	%	Municipality	%
La Pintana	98	La Florida	77
El Bosque	94	Maipú	87
La Granja	96	Estación Central	88
Cerro Navia	98	Santiago Centro	65
Renca	95	Macul	86
Huechuraba	92	Cerrillos	88
Recoleta	91	Quinta Normal	90
San Ramón	95	Independencia	86
Lo Espejo	96	La Cisterna	83
Quilicura	89	San Miguel	73
Lo Prado	93	Lo Barnechea	80
Peñalolén	88	Las Condes	42
San Joaquín	93	Ñuñoa	60
Conchalí	95	La Reina	63
Pedro Aguirre Cerda	92	Providencia	54
Pudahuel	90	Vitacura	37

Source: CASEN 2006

Segregation indices, using municipalities as geographical units are shown in Table 9. The first striking thing is the high values of dissimilarity and Isolation (the self exposure index) indices. Both are similar to values of racial segregation. Therefore, it is possible to say that Santiago is a segregated city in economic terms. The index of dissimilarity indicates that, on average, during the last 20 years, 52% of the low educational achievement households need to move from their census tracts to others with a lower proportion of them to recover the evenness (evenness means that every census tract must have the same proportion of both groups as the whole city). As mentioned, the Isolation index is high as well, and its complement, exposure, is very low.[6] For instance, Kaplan et al. (2004) show that the average isolation index in US based on race is 0.58. Although highly correlated, evenness and exposure have different interpretations. The exposure index measures the extent a individual that belongs to a minority group can interact with another group individuals, or the contact probability of people of different economic background. Therefore, in Santiago families with a head of households with a low educational achievement, have a probability of almost 90% of interacting with families of the same background. Nevertheless, the indices changes that can be observed from 1992 to 2006 are significant. The dissimilarity index shows a reduction of 32% between 1992 and 2006. The Isolation index falls 3.3% in the same period, and exposure exhibits and increment of 33%. Hence, in Santiago between 1992 and 2006 **RS** has been reduced. However, still it presents high levels of segregation.

To have a clearer picture about Santiago segregation Table 10 shows the dissimilarity index for 6 different cities: Stockholm, Lima, Mexico City, Montevideo, Bogotá and Buenos Aires. Stockholm was chosen for comparing Santiago with an egalitarian city, meanwhile the remaining cities are a group of the main Latin American capitals. Interesting is the case of Buenos Aires, Montevideo and Bogotá, because they are capitals of countries with an income per person similar to the Chilean one. Given these indices values, one can see that after the 2006 improvement of Santiago indices, they are still

[6] The summation of Isolation and Exposure is 1 only when there are two groups under analysis.

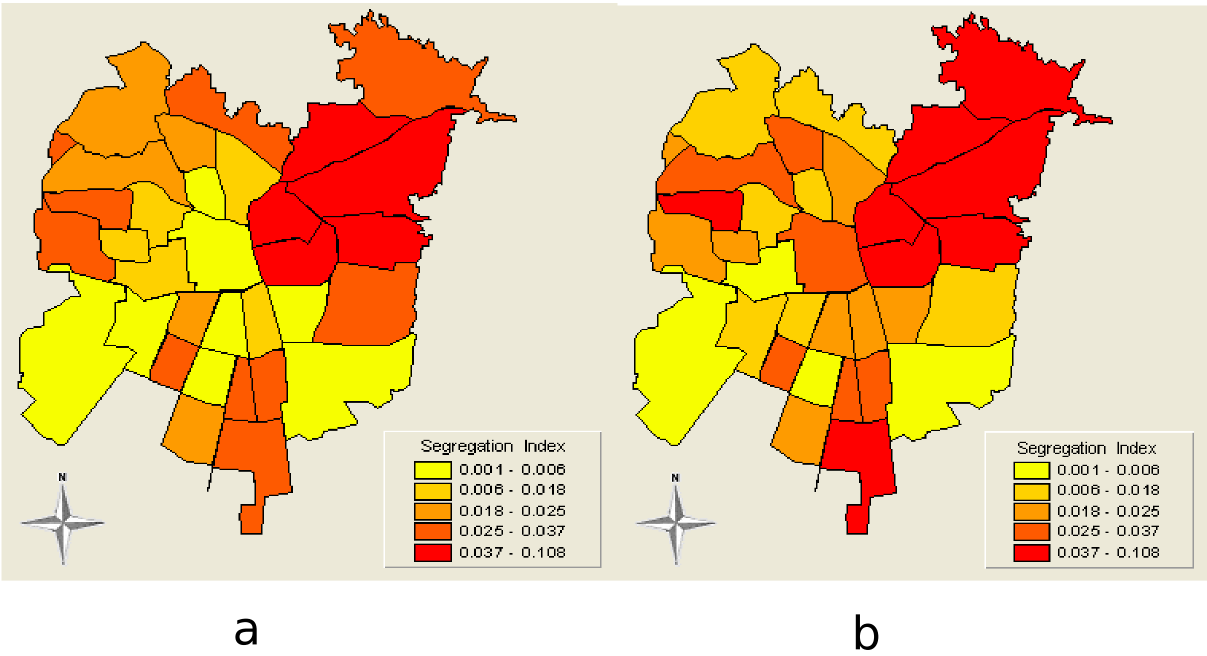

Figure 6: Santiago Municipalities level of exposure (CASEN 1992 and CASEN 2003)

higher than Buenos Aires and Montevideo ones, but lower than those observed in other capitals of Latin American countries.

Finally, it is important to bear in mind that due to the MAUP problem, the comparison is not as accurate as it could be. However, as it is indicated in the sources of these calculations, the indices were obtained using as sub-geographical units census tracts, neighbourhoods and parishes, in the case of Stockholm. In the case of Santiago municipalities were used, hence the level the segregation reported is lower than the one that would be obtained using smaller geographical units. Consequently, the differences between Santiago segregation and the one of those cities in Table 10 may be even higher. The reason for using municipalities as geographical units to calculate the segregation indices is that following this procedure it is possible to analyze Santiago **RS** evolution.

Table 9: Santiago Segregation Indices

Index	1992	1994	1996	1998	2000	2003	2006	2009
Dissimilarity	0.55	0.48	0.49	0.48	0.43	0.43	0.41	0.37
Isolation	0.91	0.89	0.89	0.88	0.89	0.86	0.87	0.88
Exposure	0.09	0.11	0.11	0.12	0.11	0.14	0.13	0.12

Table 10: Other cities Dissimilarity Index

Stockholm* (2001)	Lima** (1993)	Mexico City** (2000)	Montevideo*** (1998)	Bogotá † (2003)	Buenos Aires †† (2001)
0.214	0.42	0.38	0.35	0.47	0.268

Source:* Soderstrom and Uusitalo (2010) **Rodríguez and Arriagada (2004) ***Kaztman and Retamoso (2005) † Almonacid (2009)

†† Groisman and Suárez (2006)

5 Causes of RS

Forces driving **RS** in Santiago are explored in this section. To achieve this aim a simple methodology, based on Bayer et al. (2004) was implemented. It consists of regressing a segregation index against a set of individual characteristics. As Bayer et al. (2004) indicate, self-exposure index has the advantage, unlike many other segregation indices, that allows for the analysis of the households' propensity of any pair of socio economic background to live together, and to consider the factors that influence this propensity separately for different pairs of socio economic background. Because of this reason, the self-exposure index, or isolation index, was chosen in order to analyze the forces behind **RS**. Following this approach, the joint distribution of households and neighbourhoods at a census tract level provided by the rich micro-database used is harvested.

The main differences between this research methodology and Bayer et al. (2004) are the following. First, in Bayer et al. (2004) a linear regression is conducted. Although the reasons for doing that follows from the index structure, it has important handicaps. One problem arises because the dependent variable is a fractional response one; therefore, it is bounded between 0 and 1, and hence, the effect of any independent variable cannot be constant throughout the range of the complete set of independent variables. Another trouble is that the predicted values of a OLS regression cannot be guaranteed to lie in the range between 0 and 1. Finally, as the case of linear probability models for binary data, there is heteroskedasticity. In order to face these drawbacks, the estimation procedure developed by Papke and Wooldridge (1996) is used. This procedure corrects the problems previously indicated whereby the estimation can be conducted even in those cases where the dependent variable takes on the values 0 or 1. Besides, the conditional expectation can be obtained in a straightforward way.

A second difference with Bayer et al. (2004) is that the segregation studied here is based on economic aspects, instead of racial. This fact opens the possibility of testing new hypothesis because this sort of segregation has not been investigated with this methodology before.

As the approach proposed by Papke and Wooldridge (1996) is based on Generalized linear models, a brief explanation of this methodology is given in the next subsection.

5.1 GLM methodology

Generalized linear models (GLM hereinafter), due to Nelder and Wedderburn (1972), are flexible generalizations of ordinary least squares models. They build a connection between the dependent variable distribution and the linear predictors using a function called the link function. The GLM approach happens to be useful because it provides a general theoretical framework for many commonly encountered statistical models, and it simplifies the implementation of these different models in statistical software, since essentially, the same algorithm can be used for estimation, inference and assessing model adequacy. In other words, as Newson (2001) points out, by choice of link and variance functions (and/or transformation of the outcome variable), the user can estimate parameters that may be proportions, rates, probabilities, odds, probits or arithmetic, geometric, harmonic or algebraic means, and their differences or ratios.

McCullagh and Nelder (1989) provide a detailed survey of this sort of models and what follows is based on this work.

The traditional framework of linear regression models is : $y_i = \mathbf{x_i}\beta + \varepsilon_i$, where y_i is the dependent variable, $\mathbf{x_i}$ is a vector of independent variables, β is a vector of parameters and ε_i are the disturbances. It is assumed that $y_i \sim N(\mu_i, \sigma^2)$, where $E(y_i) = \mu_i$ and σ^2 is homocedastic variance. The latter is the stochastic component. The linear predictor $\eta_i = \mathbf{x_i}\beta$ corresponds to the systematic component. The linear predictor is a function of the mean μ_i via a link function $G(\mu_i)$. This function G for the normal linear models is an identity.

GLM models are the result of two extension of the normal linear model: firstly, stochastic components are allowed to follow distributions other than the normal and, secondly, link functions can be other than the identity. Regarding the stochastic assumptions, GLM models can be used to model variables following distributions in the exponential family:

$$f(y;\theta,\psi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\psi)} + c(y,\psi)\right\}$$

or

$$\log f(y;\theta,\psi) = \frac{y\theta - b(\theta)}{a(\psi)} + c(y,\psi)$$

where ψ is a dispersion parameter. For example, the normal distribution:

$$f(y;\theta,\psi) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{\frac{-(y-\mu)^2}{2\sigma^2}\right\}$$

with log density

$$\log f(y;\theta,\psi) = \frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\left(\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2)\right)$$

results with $\theta = \mu$, $\psi = \sigma^2$, $a(\psi) = \psi$, $b(\theta) = \frac{\theta^2}{2}$ and $c(y;\psi) = -\frac{1}{2}(\frac{y^2}{\sigma^2} + \log(2\pi\sigma^2))$. Many other distributions are in the exponential family. A collection of the most used one are shown in table 11.

Regarding link functions, they can be, in theory, any monotonic, differentiable functions. Nonetheless, in practice, only a small set of link functions can be used. In particular, links function are chosen such that the inverse link $\mu_i = G^{-1}(\eta_i)$ is easily computed, and so that G^{-1} maps from $\mathbf{x_i}\beta = \eta_i \in \Re$ into the set of admissible values for μ_i. Table 11 shows the canonical links: for instance, a log link is usually used for the Poisson model, since while $\eta_i = \mathbf{x_i}\beta \in \Re$, because y_i is a count, we have $\mu_i \in 0, 1, ..., \infty$. For binomial data, the link function maps from $0 < \pi < 1$ to $\eta_i \in \Re$, and three links are commonly used: logit:$\eta_i = \log(\pi/(1-\pi))$; probit:$\eta_i = \psi^{-1}(\mu_i)$, where $\psi()$ is the normal cumulative distribution function; complementary log-log: $\eta_i = \log(-\log(1-\mu_i))$. Note that binary data are handled in the GLM framework as special cases of binomial data.

In order to estimate the model Nelder and Wedderburn (1972) propose the **quasi-likelihood** approach. Apart from the vector of mean values μ and the matrix of covariances among the observations V, quasi-likelihood models require a score function inspired from the first order of conditions from Generalized Least Squares. The quasi-score function is:

$$D'W(y-\mu) = 0$$

where $W = V^{-1}$ and D' is a matrix of partial derivatives showing the impact of each coefficient on the predicted value:

$$\begin{bmatrix} \frac{\partial\mu_1}{\partial b_1} & \frac{\partial\mu_2}{\partial b_1} & \cdots & \frac{\partial\mu_N}{\partial b_1} \\ \frac{\partial\mu_1}{\partial b_2} & \frac{\partial\mu_2}{\partial b_2} & \cdots & \frac{\partial\mu_N}{\partial b_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\mu_1}{\partial b_p} & \frac{\partial\mu_2}{\partial b_p} & \cdots & \frac{\partial\mu_N}{\partial b_p} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{12} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \cdots & w_{NN} \end{bmatrix} \begin{bmatrix} y_1-\mu_1 \\ y_2-\mu_2 \\ \vdots \\ y_N-\mu_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

In a quasi-likelihood framework, one must begin with an estimate of the coefficients $\mathbf{b}$ and then iteratively calculate values for $\hat{\mu}$ and $\hat{W}$ and $\hat{D}$ and when the number smashing stops, one has obtained a quasi-likelihood estimator. Liang and Zeger (1986) pioneered this. They claim that the parameter estimate $\mathbf{b}$ are consistent and have many of the same properties of maximum likelihood.

For quasi-likelihood estimation and inference the precise response distribution is not specified, but rather, there is only a link function and the form of the variance function as it depends on the mean. Since quasi-likelihood estimation uses formally identical techniques to those for the Gaussian distribution, this family provides a way of fitting Gaussian models with non-standard link functions or variance functions, incidentally. This is the main difference that exists between quasi-likelihood models and maximum-likelihood models where it is assumed that there is a probability model for the data. This means that it is required to have specified a data generation mechanism, for example that the data consists of counts of events in a Poisson process. In order to propose such a mechanism, a knowledge of the physical processes that lead to the data is needed, or a substantial experience with similar data from previous studies. However, sometimes there is insufficient information about the data to specify a model for them. Under such circumstances is useful to develop analyses based on approximations to the likelihood, as quasi-likelihood model does.

5.2 GLM regressions results

As mentioned in 4.1, an alternative and convenient way to measure exposure is to calculate equation 3. In this equation the variable R_k^i is the fraction of households with economic background k living in household i's neighbourhood. It is important to note that when the dummy variable r_i takes on the value 1, if the household i is of a low economic-background and k is the households of the low income group, then 4.1 measures the self-exposure, or, in other words, the isolation. If the variable R_k^i is used as dependent variable and regressed against a determined set of head-of-household characteristics the resulting parameters will describe how these characteristics can affect the propensity of a household of low income to live with other households of low income, or what actually has actually been used here, the propensity of a household whose head has a low educational level to neighbour other households that have a head with a low educational level too.

The functional form used is the following:

Table 11: Distributions in the Exponential Family used with GLM

	Normal	Poisson	Binomial	Gamma
Notation	$N(\mu, \sigma^2)$	$P(\mu)$	$B(n, \pi)/n$	$G(\mu, \nu)$
log-density	$\frac{1}{\sigma^2}(y\mu - \frac{1}{2}\sigma^2 - \frac{1}{2}y^2) - \frac{1}{2}\log(2\pi\sigma^2)$	$y \log \mu - \mu - \log(y!)$	$n[y \log(\frac{\pi}{1-\pi}) + \log(1-\pi) + \log \binom{n}{ny}]$	$\nu(-\frac{y}{\mu} - \log \mu) + \nu \log y + \nu \log \nu - \log \Gamma(\nu)$
Range of y	$(-\infty, \infty)$	$0, 1, ..., \infty$	$\frac{z}{n}, z \in \{0, 1, ..., n\}$	$(0, \infty)$
ψ	σ^2	1	n^{-1}	ν^{-1}
$b(\theta)$	$\frac{\theta^2}{2}$	$\exp(\theta)$	$\log(1 + e^\theta)$	$-\log(-\theta)$
$c(y; \psi)$	$-\frac{1}{2}\left(\frac{y^2}{\psi} + \log(2\pi\psi)\right)$	$-\log(y!)$	$\log \binom{n}{ny}$	$\nu \log(\nu y) - \log y - \log \Gamma(\nu)$
$\mu(\theta) = E(y; \theta)$	θ	$\exp(\theta)$	$\frac{e^\theta}{1+e^\theta}$	$\frac{-1}{\theta}$
Link, $\theta(\mu)$	$\theta = \mu$	$\log(\mu)$	$\log\left(\frac{\pi}{1-\pi}\right)$	μ^{-1}
Variance function $V(\mu)$	1	μ	$\mu(1-\mu)$	μ^2
Deviance	$\sum(y_i - \hat{\mu_i})^2$	$2\sum[y_i \log\left(\frac{y_i}{\hat{\mu_i}}\right) - y_i + \hat{\mu_i}]$	$2\sum[y_i \log\left(\frac{y_i}{\hat{\mu_i}}\right) + (n_i - y_i)\log\left(\frac{n_i - y_i}{n_i - \hat{\mu_i}}\right)$	$2\sum[-\log\left(\frac{y_i}{\mu_i}\right) + \frac{y_i - \hat{\mu_i}}{\hat{\mu_i}}]$

$$E(R_k^i|r_i x_i) = G(r_i x_m^i \beta) \quad (4)$$

where β is a vector of coefficients explaining how attribute x_m affects the isolation. McCullagh and Nelder (1989) suggest to use the logistic function as the link function, hence $G(z) \equiv \Lambda(z) \equiv \frac{\exp(z)}{1+\exp(z)}$.

The estimation made is a quasi-likelihood, using a Bernoulli log-likelihood function given by:

$$l_i(\mathbf{b}) \equiv R_i \log [G(r_i x_i \mathbf{b})] + (1 - R_i) \log [1 - G(r_i x_i \mathbf{b})] \quad (5)$$

The reasons for choosing the Bernoulli log-likelihood function, as Papke and Wooldridge (1996) indicate, are as follows. First, the maximizing process of this function is easy. Second, as the Bernoulli log-likelihood function belongs to the linear exponential family, the quasi-maximum likelihood estimator, $\hat{\beta}$, is consistent for β as long as the equation 4 holds. In other words, $\hat{\beta}$ is consistent and $\sqrt{N}$- is asymptotically normal regardless of the distribution of y_i is conditional to x_i; y_i could be a continuous variable, a discrete variable, or have both continuous and discrete characteristics. In some cases for fractional data, $\hat{\beta}$ is efficient in a class of estimators containing all quasi-maximum likelihood estimators in the linear exponential family and weighted non-linear estimators.

Although in trying to explain **RS**, based on income it is natural to think that income itself has to be a fundamental driving force, working through households' budget constraints that influence residential choices. However, following **?** arguments, it would be possible to find a number of alternative mechanisms only indirectly related to income that may drive **RS**, and an important amount of **RS** can be explained by sorting on the basis of these other mechanisms, particularly if the correlation among **RS** and these other households characteristics is strong.

The aim of the present research is to examine the extent to which cross-education differences in head-of-households characteristics can explain the observed pattern of **RS** based on income. As mentioned above, educational achievement was used because it is a better indicator of permanent income. Table 12 shows the variable definitions.

Table 12: Independent variables description

Variables	**Description**
bad health	a dummy variable assuming value 1 if the household head has been in the hospital during the last 6 months
divorced	a dummy variable assuming value 1 if the household head is divorced
dwelling rent	amount of money paid as rent, in the case of ownership it is an imputed rent
ethnicity	a dummy variable assuming value 1 if the household head belongs to one of the native ethnic group
female	a dummy variable assuming value 1 if the household head is a woman
household members	number of household members
illiteracy	a dummy variable assuming value 1 if the household head does not have the ability to read or write
income	the household income per month
joblessness	a dummy variable assuming value 1 if the household head does not have a job
not married partners	a dummy variable assuming value 1 if the household head is not officially married to his/her partner
over 60	a dummy variable assuming value 1 if the household head is over 60 years old
ownership	a dummy variable assuming value 1 if the household head is the owner of the house where the household lives
single	a dummy variable assuming value 1 if the household head is single
subsidy	a dummy variable assuming value 1 if the household has been recipient of the affordable housing subsidy
widow	a dummy variable assuming value 1 if the household head is a widow

Seven regressions were performed for years 1992, 1994, 1996, 1998, 2000, 2003 and 2006. The results of these estimations are shown in Table 13. Columns show the marginal effects. Asterisks indicate that the variable is significant at %5.

As can be seen, the affordable housing subsidy is always positive and significant. If the variables that have always been in the survey are considered, it is the second most important segregation driver, after ownership, with an average impact of 0.029. Besides, this variable impact has been growing, particularly in 2003 and 2006. The reason of this result relies on evidence of the way Chilean public housing policy has worked, particularly during the nineties and the beginning of the present decade. As a way to reduce the lack of housing in poorer sectors of society, the Chilean Government implemented an aggressive policy to build affordable dwellings. In order to assure affordability and profitability, these dwellings were made using cheap land, typically located far away from the city centre, mainly in the South end of Santiago. It has been claimed that this mechanism generates segregation (Sabatini et al., 2001) but this conjecture has never been tested before.

The first force explaining **RS** in Santiago is ownership, which has, on average, an impact of 0.045. This variable, however, is related to subsidies, because, given that subsidies are given for buying and not for renting, if a households has received the affordable housing subsidy, it must be the owner of the dwelling. Consequently, these two variables explain the impact of the housing policy on segregation. The third more important segregation driver is illiteracy, in the seven regressions it is significant and has a positive impact on segregation. This variable is highly related to work quality, unemployment and permanent income. Households number is also a segregation driver. It has an average impact of 0.012. The dwellings' rents are always significant as well and have a negative impact on segregation, hence the higher the rent the lower the low-income-households' propensity of neighbouring other households of the same socio economic condition. Another important segregation driver has been ethnicity. However, that question has not always been addressed by the survey, therefore it is possible to test the hypothesis' importance only for years 1996, 2000, 2003 and 2006. In these years this variable has had a positive and significant effect on segregation. In fact, in 2006, it is the most important

Table 13: Marginal effects of glm regressions 1992-2006

Variable	1992	1994	1996	1998	2000	2003	2006
subsidy	0.024*	0.026*	0.025*	0.019*	0.018*	0.042*	0.044*
ethnicity	-	-	0.013*	-	0.021*	0.026*	0.124*
divorced	0.004	0.012*	0.017*	-0.001	0.016*	0.023*	-0.015
ownership	0.034*	0.019*	0.034*	0.068*	0.062*	0.034*	0.066*
widow	0.000	0.020*	0.010*	0.004	0.005	0.002	-0.012
bad health	0.010*	0 .010*	0.014*	0.011*	0.014*	0.010	0.004
illiteracy	0.019*	0.023*	0.021*	0.036*	0.022*	0.049*	0.028*
joblessness	0.006	0.010	0.016*	0.017*	0.007	0.018*	0.016
dwelling rent	-6.83e-07*	-2.72e-07*	-5.39e-07*	-7.04e-07*	-1.86e-07*	-4.87e-07*	-1.62e-06*
female	0.005	-0.023*	-0.010*	-0.015*	-0.018*	0.002	0.027*
over 60	0.001	-0.001	-0.004	-0.005	-0.005	-0.016*	-0.001
household members	0.010*	0.010*	0.008*	0.013*	0.007*	0.021*	0.014*
N	9,874	7,531	7,254	10,502	9,825	8,606	8,433
AIC	0.356	0.435	0.434	0.492	0.442	0.539	0.722
BIC	-90,171	-66,693	-63,992	-96,429	-89,661	-77,192	-74,930

segregation driver. This result is quite intuitive and it follows what literature has pointed out concerning the relationship between ethnicity and low-income.

6 Final Remarks

Chilean Public Housing Policy has been based on giving subsidies to low income households in order to facilitate their access to affordable dwellings. Given the momentous lack of this kind of housing units, during the nineties, this policy was very aggressive, and more than 400,000 units were built. Almost 70% of these units were built in new neighbourhoods located in the urban fringe giving very limited degree of freedom to households in order to choose among them. Although this policy has been very successful in both reducing the lack of affordable dwellings and through this process redistributing income, the fact that the most of the beneficiary households have been located in the city boundary has been the main argument for saying this policy is one of the most important **RS** drivers in Santiago. Despite the need for elucidating whether the argument given above is true, there was not an attempt for testing this hypothesis before. This research fills this gap. An econometric model was developed to find out if the housing policy is a **RS** driver.

The regressions results indicate that subsidies given by the state for buying a social dwelling are a significant determinant of **RS**.

This implies, in fact, that housing policy has exogenously raised **RS** by the mechanism of buying cheap land far from the city centre as locations for building social dwellings. As a matter of fact, the average impact of subsidies on isolation is 0.029, which means that if a low-income household has received affordable housing subsidies, the propensity of neighbouring households of the same socio economic condition will increase by 2.9%.

One of the main reasons is the lack of a competitive private market for affordable dwellings. This lack of development has limited the offering of vouchers as a direct subsidy to the demand instead of a dwellings' provision. Therefore, the policy efforts should be on trying to establish the conditions for a competitive market. Despite the fact that since

2000 some changes to the housing policy have been implemented by the Chilean State, the positive effects of the housing policy on segregation has been increasing.

Regarding the proportion of households with a head of a low level of education, one can observe a reduction between 1992 and 2006, something that may be correlated with the economic growth of the country in the same period.

An important fact is that dissimilarity, isolation and exposure indices in this period have exhibited important reductions, although they are still high compared to other countries with a similar income per person. This could indicate that although the country has become richer, the difference between well-off people and poor people has remained constant, and this difference has a spatial representation. The latter is only a conjecture and further research is needed in order to test its validity.

A significant amount of **RS** is explained by a set of other households' attributes, like illiteracy, ethnicity and households members, giving empirical evidence to the ideas developed by Schelling.

The methodology used here is a simple way to intuitively characterize **RS**, also taking into account how to deal with fractional response dependent variables. Finally, it is important to say that this is the first time that a model of this sort has been used to study economic **RS**.

References

Abramson, A., Tobin, M. and Vander Goot, M.: 1995, Geography of Metropolitan Opportunity: The Segregation of the Poor in US Metropilitan Areas, *Housing Policy Debate* **6**, 45–72.

Almonacid, J.: 2009, Análisis espacial de la segregación residencial en bogotá. años 1993 y 2005, *12 Encuentro de Geografos de Latino América*.

Bayer, P., McMillan, R. and Rueben, K.: 2004, What drives racial segregation, *Journal of Urban Economics* **56**, 514–535.

Burnell, J.: 1988, Crime and racial composition in contiguous communities as negative externalities: Prejudiced households' evaluation of crime rate and segregation nearaby reduces housing values and tax revenues, *Journal of Economics and Sociology* **47(2)**, 177–193.

Burton, E.: 2003, Housing for an urban renaissance: Implications for social equity, *Housing Studies* **18(4)**, 537–562.

Castillo, M.: 2008, Beyond the quantitative results, the challenges of housing policy in Chile, *Quorum* **20**, 14–29.

Charles, C., Dinwiddie, G. and Massey, D.: 2004, The continuing consequences of segregation: Family stress and college academic performance, *Social Science Quarterly* **85(5)**, 1353–1373.

Cheshire, P.: 2007, Segregated neighbourhoods and mixed communities: a critical analysis, *Joseph Rowntree Foundation*.

Christopher, A.: 1990, Apartheid and urban segregation levels in South Africa, *Urban Studies* **27**, 421–440.

Clapp, J. and Ross, S.: 2004, Schools and housing markets: An examination of school segregation and performance in Connecticut, *Economic Journal* **114(499)**, 425–440.

Dawkins, J., Shen, Q. and Sanchez, T.: 2005, Race, space, and unemployment duration, *Journal of Urban Economics* **58**, 91–113.

Dosh, P.: 2003, Review essay: Violence, spatial segregation, and the limits of local empowerment in urban Latin America, *Latin American Politics and Society* **45(4)**, 129–146.

Duncan, O. and Duncan, B.: 1955, A methodological analysis of segregation indices, *American Sociological Review* **20**, 210–217.

Galetovic, A. and Poduje, I.: 2006, *Quién es Santiago*, Centro de Estudios Públicos, chapter 1, pp. 3–23.

Groisman, F. and Suárez, A.: 2006, Segregación residencial en la ciudad de buenos aires, *Población de Buenos Aires* **3**(4), 27–37.

Hidalgo, R.: 2002, Vivienda social y espacio urbano en Santiago de Chile. Una mirada retrospectiva a la acción del estado en las primeras décadas del siglo xx, *EURE*.

INE: 2005, *Chile: Proyecciones y Estimaciones de población. Total País: 1950-2050*, Instituto Nacional de Estadísticas de Chile.

James, D. and Taeuber, K.: 1985, Measures of segregation, *Sociological Methodology* **17**, 315–321.

Kaplan, D., Wheeler, J. and Holloway, S.: 2004, *Urban Geography*, John Wiley & Sons.

Kaztman, R. and Retamoso, A.: 2005, Spatial segregation, employment and poverty in montevideo, *CEPAL Review*.

King, T. and Mieszkowski, P.: 1973, Racial discrimination, segregation, and the price of housing, *Journal of Political Economy* **81**, 590–606.

Krugman, P.: 1996, *The Self Organizing Economy*, Blackwell.

LaVeist, T.: 2003, Racial segregation and longevity among African Americans: An individual-level analysis, *Health Services Research* **38**, 1719–1733.

Liang, K. and Zeger, S.: 1986, Longitudinal data analysis using generalized linear models, *Biometrika* **73**, 13–22.

Logan, J. and Messner, S.: 1987, Racial residential segregation and suburban violent crime, *Social Science Quarterly* **68(3)**, 510–527.

Madden, J.: 2001, Do racial composition and segregation affect economic outcomes in metropolitan areas?, *Problem of the century: Racial stratification in the United States*, Russell Sage Foundation.

Massey, D.: 2001, Segregation and violent crime in urban America, *Problem of the century: Racial stratification in the United States*, Russell Sage Foundation.

Massey, D. and Denton, N.: 1988, The dimensions of residential segregation, *Social Forces* **67**, 281–315.

Mattos, C.: 2002, Santiago de Chile faces globalisation: another city?, *Revista de Sociología y Política* **19**, 31–54.

McCullagh, P. and Nelder, J.: 1989, *Generalized Linear Models*, Chapman and Hall.

Nelder, J. and Wedderburn, R.: 1972, Generalized linear models, *Journal of the Royal Statistical Society* pp. 370–384.

Newson, R.: 2001, Review of generalized linear models and extensions by Hardin and Hilbe, *The Stata Journal* **1(1)**, 98–100.

OECD: 2004, Oecd economic outlook.

Papke, L. and Wooldridge, J.: 1996, Econometric methods for fractional response variables with an application to 401 (k) plan participation rates, *Journal of applied econometrics* **11**, 619–632.

Reardon, F. and Firebaugh, G.: 2002, Measures of multigroup segregation, *Sociological Methodology* **32**, 33–67.

Rodríguez, J.: 2001, Segregación residencial socioeconómica, qué es, cómo se mide, qué está pasando, importa, *Serie población y desarrollo, ECLAC*.

Rodríguez, J. and Arriagada, C.: 2004, Segregación residencial en la ciudad latinoamericana, *EURE* **XXIX**(89), 5–24.

Ruprah, I. and Marcano, L.: 2008, Chile's housing finance: A story of success?, *Housing Finance International*.

Sabatini, F., Cáceres, G. and Cerda, J.: 2001, Segregación residencial en las principales ciudades chilenas: tendencias de las tres últimas décadas y posibles cursos de acción, *EURE*.

Schelling, T.: 1971, Dynamics models of segregation, *Journal of Mathematical Sociology* **1**, 143–186.

Simian, J.: 2010, Logros y desafíos de la política habitacional en Chile, *Centro de estudios públicos*.

Soderstrom, M. and Uusitalo, R.: 2010, School choice and segregation: evidence from an admission reform, *Scandinavian Journal of Economics* **112**(1), 55–76.

Sugranyes, A.: 2004, "housing Policy in Chile: Liberal success - roof for the poor", *Habitat International Coalition Documents*.

Tokman, A.: 2006, *El MINVU, la política habitacional y la expansión excesiva de Santiago*, Centro de Estudios Públicos, chapter 17, pp. 462–89.

Wilson, F. and Hammer, R.: 2001, Ethnic residential segregation and its consequences, *Urban inequality: Evidence from four cities*, Russell Sage Foundation.

Yinger, J.: 2001, Housing discrimination and residential segregation as causes of poverty, *Understanding poverty*, Harvard University Press.

Young, H.: 1998, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princenton University Press.

Zhang, J.: 2004, Residential segregation in an all-integrationist world, *Journal of Mathematical Behavior and Organization* **54(4)**, 533–550.