# The Persistent Inequality in the Great Brazilian Cities: The Case of Brasília

Ferreira Lima, Luis Cristovao

University of Brasilia

September 2013

# THE PERSISTENT INEQUALITY IN THE GREAT BRAZILIAN CITIES: THE CASE OF BRASÍLIA

**Luis Cristovao Ferreira Lima***

**ABSTRACT**

Using the censuses of 2000 and 2010, we have noticed that the inequality of the household *per capita* income in the biggest Brazilian cities did not show a trend of reduction, differently from the whole country. Also, the inequality in those cities is substantially higher than the Brazilian. We investigate the determinants of this high and persistent inequality for Brasília (Federal District). We use the static decomposition of the generalized entropy indexes and the decomposition by regression with the method of Fields and the Shapley value. We verified that the public sector was the main factor to explain why the inequality was kept high in the capital of Brazil. All the methods reached the same conclusion. While the shrinking differences on the education attainment of the population had an effect of reducing the inequality, the policy of paying better salaries to the public servants had the opposite effect, which preserved the high inequality. This policy induces the migration to Brasília and it has a long run impact on the retirements and pensions benefits, which perpetuates the disparities.

**Key-Words:** Brasília (Federal District); Public Sector; Inequality Decomposition; Generalized Entropy Indexes; Fields method; Shapley value.

***JEL Classification:*** *C21, C71, D31, I24, J31, O15.*

* University of Brasília, Brasília – Brazil. Email for contact tovaodf@hotmail.com. Paper presented as Bachelor's Thesis at the University of Brasília, August 2013. Supervisor: PhD Michael Christian Lehmann.

# 1. Introduction

The current fall on Brazilian inequality has been significantly documented and studied. After 40 years of increasing disparities, the inequality has finally started to diminish consistently during the first decade of the twenty-first century (. A higher minimum wage, the conditional cash transfer programs such as *Bolsa Família*, the universalization of the social security system, the greater schooling of the population and the shrinking regional disparities have been the main determinants for this reduction (Ferreira et al, 2006; Barros et al, 2007; Souza and Osório, 2011; Azzoni and Neto, 2011).

In this context, it is pertinent to study how inequality has behaved in a disaggregated level. There are studies showing smaller differences between Brazilian regions (Ferreira et al, 2006; Azzoni and Neto, 2011; Santos and Ferreira, 2007) and also convergence between Metropolitan Areas and other areas of the country (Souza and Osório, 2011). Souza and Osório (2011) have shown that, from 1981 to 2009, there was a drop on the difference of the household per capita income in Metropolitan Areas on the rest of the country. The Metropolitan Areas used to be much richer than other areas, but this pattern has changed. This has been an important component of the falling inequality (or a consequence of it), especially on the 2000's. The authors concluded that Metropolitan Areas have lost its relative dynamics in comparison to other areas, as long as the labor income has not grown much. Also, there was a process called of industry spread, which favors the rest of the country instead of big cities. Facing this, it is interesting to investigate the trend of inequality on the great Brazilian cities. If there is a relative deterioration on the conditions of those cities, how has inequality behaved inside them? Have these cities followed the same trend of equalization that Brazil as a whole did? This paper tries to answer those questions to the case of Brasilia (Federal District), capital of Brazil and fourth biggest city in the country.

The investigation of inequality in a disaggregated level is important, as long as the perception of people is directly related to what they see around them. Even with a considerable fall on the Brazilian inequality, if it has not happened on the big cities, their habitants would not perceive this drop (and they are 50% of the population). The idea of lower inequality would be vague and not verified on the environment those people live.

Using the Brazilian Censuses of 2000 and 2010, the table 1 below shows the different trends of inequality in Brazil as a whole and in the big cities. We use two types of income, the household per capita income and wages. We must make two important notes: the

inequality measured using the Census is higher than the one measured using the National Survey of Households (*Pesquisa Nacional por Amostra de Domicílios* – PNAD). The latter survey is conducted every year and it uses a smaller sample. As the sample of the Census is bigger, it is possible to capture more the extremely rich people and their income (Souza, 2013). On the report of the Census 2010, made by the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística* - IBGE), the inequality reported has been computed using only wages (IBGE, 2011). When we measure the inequality of the household per capita income, it is higher and with a less significant tendency of reduction. This can also be seen on the data compiled by DATASUS (2013).

**Table 1 – Gini Index of the ten most populous Brazilian cities and the interior of their States, 2000 and 2010**

| | **Household Per capita income** | | **Wages** | |
|---|---|---|---|---|
| **City** | **2000** | **2010** | **2000** | **2010** |
| São Paulo | 0.616 | 0.626 | 0.573 | 0.557 |
| Rio de Janeiro | 0.614 | 0.623 | 0.579 | 0.556 |
| Salvador | 0.652 | 0.630 | 0.612 | 0.557 |
| Brasília | 0.627 | 0.628 | 0.604 | 0.595 |
| Belo Horizonte | 0.617 | 0.606 | 0.598 | 0.559 |
| Fortaleza | 0.648 | 0.617 | 0.605 | 0.554 |
| Manaus | 0.638 | 0.611 | 0.570 | 0.523 |
| Curitiba | 0.593 | 0.557 | 0.581 | 0.521 |
| Recife | 0.676 | 0.682 | 0.637 | 0.595 |
| Porto Alegre | 0.604 | 0.606 | 0.585 | 0.557 |
| **Brazil** | **0.645** | **0.596** | **0.596** | **0.526** |
| | | | | |
| | **Household per capita income of the State, excluding the capital** | | **Wages of the State, excluding the capital** | |
| **State** | **2000** | **2010** | **2000** | **2010** |
| São Paulo | 0.566 | 0.515 | 0.537 | 0.467 |
| Rio de Janeiro | 0.568 | 0.524 | 0.516 | 0.457 |
| Bahia | 0.624 | 0.572 | 0.569 | 0.515 |
| Distrito Federal | 0.627 | 0.628 | 0.604 | 0.595 |
| Minas Gerais | 0.592 | 0.520 | 0.562 | 0.466 |
| Ceará | 0.611 | 0.537 | 0.571 | 0.483 |
| Amazonas | 0.653 | 0.591 | 0.542 | 0.513 |
| Paraná | 0.580 | 0.498 | 0.563 | 0.470 |
| Pernambuco | 0.626 | 0.555 | 0.563 | 0.480 |
| Rio Grande do Sul | 0.555 | 0.501 | 0.539 | 0.463 |

Source: Censuses 2000 and 2010, IBGE. Author's calculations.
Note 1: On the computation of the indexes, we have excluded all the null incomes. It can underestimate the inequality in the year 2010, as long as there were null incomes this year but not in 2000. On the next sections, the inequality measured for Brasilia will use some null incomes, making it higher than shown above.
Nota 2: We have used the municipality, not the Metropolitan Areas. The indexes for the Metropolitan Areas are very similar and show the same pattern. They can be seen on DATASUS (2013).

The table shows the Gini Index of the household per capita income and the wages on the ten biggest Brazilian municipalities (all of them are State capitals). The table also shows the index for the rest of the State, excluding its capital. We notice a trend of fall on the inequality of wages on the big cities and also on the interiors. This fall, however, is stronger on the interior of the States. Also, the inequality of the household per capita income drops on the interiors, but it does not have a uniform pattern for the populous cities. Salvador, Belo Horizonte, Fortaleza, Manaus and Curitiba have had a reduction on inequality. On the other hand, São Paulo, Rio de Janeiro, Brasília, Recife and Porto Alegre have had an elevation. All of them have shown a small variation, giving weak indications of changes between 2000 and 2010. Roughly, the level of inequality has kept almost the same after 10 years.

The results on table 1 show the inequality on the most populous cities is greater and does not have an observable trend. The results for the interiors of the States suggest the fall on inequality was more concentrated on municipalities of small or medium size. The data emphasizes the heterogeneous pattern of inequality dynamics in Brazil. The habitants of big Brazilian cities live with an inequality only inferior to the big African cities (United Nations Habitat, 2008). Out of the ten most populous Brazilian cities, only Curitiba has a Gini index of the household per capita income lower than 0.6. Concerning the States after excluding the capital, all of them present a Gini Index lower than 0.6.

The investigation of why inequality is not falling on big cities is a relevant issue for new studies. An elevation on inequality of the household per capita income together with a drop on wages inequality can be justified by the following: change on the configuration of the families' conditional to their income or rise on the inequality of the income from other sources. Analyzing the former, the size of richer families must drop more than the size of poorer ones. Or, the number of workers on the richer families have to increase more than on poorer ones. It could happen through marriage of people of the same income class and the insertion of siblings on the labor market. Analyzing the latter, the income from other sources,

such as rentals, pensions, interests and direct transferences has to be more concentrated. From 2000 to 2010, the direct transference of money has been intensified (e. g. *Bolsa Família*). Hence, the inequality of the income from other sources has to compensate this equalizing effect of the transferences by a greater concentration of pensions and property (wealth).

The data suggests each city has particularities, as long as the inequality has shown different patterns. Understanding this can help on the development of regional policies and in a more equitable growth. On this context, it is important to know the migration of those cities, its labor market, the configuration of their families and the concentration of income from other sources. This paper investigates the particularities of Brasilia (Federal District) on this context of urban inequality in Brazil. As it is going to be shown on the later, the inequality in Brasilia has increased slightly between 2000 and 2010, with a great influence of the public sector. The income of the public servants has increased much more than the general population, which has brought to more inequality.

In the Federal District, the inequality of the household per capita income and wages has kept very high. Among the ten biggest Brazilian cities, Brasilia has shown the smallest reduction on the wages inequality. Analyzing the particularities of Brasilia, we can see that the migration and the income from other sources have a role on this high inequality. Although, the most relevant to explain the inequality in Brasilia is the idiosyncratic labor market of the capital of the Republic. The effects of the public sector are essential to explain the preservation of the high levels of inequality in the Federal District. Between 2000 and 2010, the wages of the public servants have increased much more than the average. By itself, this elevation creates more disparities, induces more qualified workers to opt for the public sector and attracts new immigrants (Souza and Medeiros, 2013a; Holanda, 2009; Da Mata et al, 2007). Also, this also causes more concentration of pensions, as the public sector gives benefits much more generous to inactive workers than private companies do (Souza e Medeiros, 2013b). This study is based on the solid data of the Brazilian Censuses of 2000 and 2010, as well as modern techniques of inequality decomposition to demonstrate the contribution of different individual characteristics on the income distribution on the two years under analyses. These techniques can isolate the effects of the public sector remuneration, and the results show that this was the main cause for the maintenance of the high inequality in the Federal District of Brazil.

In order to access the determinants of inequality in 2000 and 2010, we make a static decomposition and also a regression-based decomposition using the Fields' Method and the Shapley value, the strongest technique for this kind of analyses. The decomposition using the

Shapley value has never been applied to Brazilian data. The great strength of the regression-based decomposition is its capacity to isolate the effects of each variable on total inequality, as a regression. This decomposition consists of estimating an income generating function and decomposes the result by any inequality index (Wan, 2004). Hence, we can identify the contribution of each individual characteristic on total inequality. The education had a trend to reduce inequality, and, *ceteris paribus*, would make inequality to fall. Nevertheless, there is an elevation on the inequality related to the type of occupation of a person, basically related to the remuneration of public servants. All the methods employed have reached the same conclusions.

After this introduction, there is a section discussing the possible causes of the inequality on the Federal District and its trend between 2000 and 2010, emphasizing the wages of the public sector. Section 3 shows the methodologies used: the static decomposition and the regression-based decomposition of inequality. Section 4 shows the results for the household per capita income using the two methods. Section 5 concludes. At the end of the text, there are two appendixes, one with the description of the database used and other with the regression based decomposition applied to wages of Brasilia's formal labor market.

## 2. Determinants of inequality – Brasilia, Federal District of Brazil

The inequality in a specific period of time has a plenty of determinants associated to personal characteristics. As we are dealing with the household per capita income (it is the same for every member of the household/family), we use the household head characteristics in order to access their contribution to inequality. Between 2000 and 2010, it is expected that the proportion of inequality attributed to each characteristic has changed, which point out new social features and also policies implemented before.

It is reasonable that a great part of the inequality is generated by differences of individual attributes. Education would be the most important of them. A more educated person will have higher income, *ceteris paribus*. In Brazil, as well as in Brasilia, between 2000 and 2010, education had a trend of reducing inequality. The returns to education have shrunk as a consequence of the universalization of the primary education and more people with higher education (Ferreira et al, 2006; Barros et al, 2007). The Federal District has also followed this trend. Another attribute is the age / experience of a person. Older people tend to have higher income, as they have accumulated more properties and have more experience on the labor market. Nevertheless, it is not clear whether the impact of this characteristic has changed on the first decade of the century.

Gender and race/color differences also play a role on inequality, given that minorities have limited access to education and can suffer from discrimination. However, those differences seem to be reducing (Ferreira et al, 2006). Hence, we ought to observe a smaller inequality coming from those groups. There is also inequality between geographical regions. People with the same attributes but living in different places can systematically show income disparities. In Brazil, those differences have fallen (Souza and Osório, 2011; Azzoni and Neto, 2011). Brasilia (Federal District) is divided into Administrative Regions and there are significant differences between them. It is impossible to affirm if there is a fall on the regional disparities of the city. It is visible that really poor and really rich neighborhoods still live side by side on the Federal District.

The family size can also affect the inequality of the household per capita income. Bigger families (with more children and elderly people) tend to have less disposable income for each member. Considering that families, in a country like Brazil, tend to diminish with time, if the number of habitants of poor households has a trend of smaller reduction than the

rich ones, inequality would rise, *ceteris paribus*. However, in the Federal District between 2000 and 2010, the fall on the number of habitants per household was basically the same for every quantil of the income distribution, around 14%.

Finally, another characteristic which can contribute to inequality is the type of occupation of the person. This occupation has an effect on income. Part of this effect is correlated to other attributes of the person, but another part comes from the particularities of the occupation. These particularities can be considered as institutional determinants of inequality. The public sector remunerates its employees better than the private sector, independently of the individual attributes (Bender and Fernandes, 2009; Vaz and Hoffman, 2007). Besides this, the wages of the public sector has risen more rapidly in the decade under analysis. Brasilia is the headquarters of the federal administration. Because of that, it shows the greatest proportion of public servants between all Brazilian big cities (MPOG, 2010). So, we would expect a great and crescent proportion of inequality coming from differences of types of occupation. Below, there is an explanation of how the public sector wages can affect inequality. The other subsection shows the migration of the Federal District, as changes on population can affect inequality.

### 2.1 Public sector wages: why so high and increasing?

The public employment has fundamental differences compared to the private one. The main distinction is the way that wages are determined. While firms has to deal with a problem of profit maximization / cost minimization, the employment on the public sector is influenced by political issues (Souza and Medeiros, 2013b; Bender and Fernandes, 2009; Vaz and Hoffmann, 2007; Holanda, 2009; Heitmueller, 2006). The government can use the wages policy to enhance its popularity (Heitmuellller, 2006). Also, the workers in the public sector are more organized and influent. In Brazil, public servants participate more in Unions which have a great bargaining power. They have also some guarantees in the law, which pressures for constant increases on wages (Souza and Medeiros, 2013a; Vaz and Hoffmann, 2007). According to Gustafsoon and Johansson (1999), countries with more workers in Unions have smaller inequality. But, if this unionization is corporatist and is present only in some groups, it would have a regressive effect. This is what seems to happen in Brazil. This pattern, *ceteris paribus*, leads to an increase on the wages differences and, consequently, an inequality

elevation (Souza and Medeiros, 2013b).

Nevertheless, as argued by Heitmueller (2006), the rules in the public sector are different from the private one, as the workers have more stability and a more generous social security program. These better conditions, hypothetically, could be the compensations for a smaller salary as a public servant comparing to a private worker. But, it is the greater power of negotiation and the government's necessity of maximizing its political capital that jointly cause more generous rules and better salaries for public servants. In Brazil, these two topics seems not be substitute, but complementary (Holanda, 2009). According to this author, public servants are responsible for goods and services, while they maximize votes and the social welfare, which could be a reason for their higher wages. But, in Brazil, as put by Vaz and Hoffmann (2007), public sector wages tend to increase with the age of the employee, and more productivity is not needed for it to occur. There is a continuous wage progressivity on the public sector, which is unimaginable on the private sector. Under these circumstances, inequality is caused by institutional specificities of the public sector.

The government, when employing, must offer a salary and conditions that attracts qualified workers. On the maximization of welfare and votes that each government has to deal, the public servants would be the main 'input'. That way, the higher salaries are justifiable in order to keep the efficiency/productivity of those employees (efficiency wages models) and inhibit corruption, as better remunerations reduces the willingness to take risks for personal benefits (Heitmueller, 2006). Although it is important to notice in which moment these wages overtake the benefit generated by them and start to cause distortions on the economy. The government should look for the smallest wage which would still attracts qualified workers to the positions opened (Holanda, 2009). Holanda (2009) and Holanda and Barbosa (2010) argue that workers choose the sector they are going to work endogenously. This would be influenced by their risk aversion, the mean remuneration of the sector and the distribution of wages inside each sector. If there is an excessive growth on the remuneration of one sector (above the equilibrium), this sector would attract more employees, which, in normal conditions would not happen. This pattern generates inefficiencies. In Brasília, the externalities caused by the public sector remuneration are gradually more visible. Between qualified workers, there is an excessive demand to this sector, in detriment to the private sector or entrepreneurship. Between 2000 and 2010, the censuses indicated an increase on the number of public servants and a drop on the number of employers among the household heads. This is, probably, a consequence of these distortions mentioned above.

There are two effects on the gross wage difference between workers in the public and private sectors: a composition effect and a segmentation effect (Souza and Medeiros, 2013b). The first one is related to differences on the qualification of the workers. The public administration demands, on average, more qualified workers than the private sector. These workers would integrate the upper tail of the income distribution, independently of the rules determining the salaries. The segmentation effect, also known as price effect, is related to the particularities of each sector. As shown above, the public sector has more generous rules and does not have to maximize profit. These characteristics make a public servant get a better remuneration than another worker with the same qualification on the private sector, on average. This phenomenon is observed in developed economies and also in the developing World (Holanda, 2009). The unbiased wage difference between the two sectors can be obtained using regression, as in Vaz and Hoffmann (2007), Bender and Fernandes (2009), Holanda and Barbosa (2010) and Souza and Medeiros (2013a).

Those researches identify a constant elevation on the wages of the public sector after *Plano Real*[1]. Bender and Fernandes (2009) use the National Household Survey (PNAD) from 1992 to 2004 in order to determine the historical pattern of the wage differential between the public and private sectors in Brazil. They found a gross differential of 64% in 1992 and 111% in 2004. The controlled differential has grown, starting at 10% in 1992 and reaching 35% in 2010. Vaz and Hoffman (2007), using the PNAD from 1992 to 2005, analyzed the services sector and got the similar results. Using the data from PNAD 2009 with an endogenous sector selection model, Holanda and Barbosa (2010) found that the hourly wage of public servants (including the employees of state companies) was considerably higher than the employees in the private sector. They have also noticed that the difference was higher for federal employees, followed by State ones, but not for municipalities' servants. Souza and Medeiros (2013a), also using the PNAD 2009, decomposed the Gini Index by income factors and observed that the wage differential corresponded to 3.1% of total inequality in Brazil (using the household per capita income).

All the studies confirm the constant growth in the difference between public sector and private sector wages in Brazil. The Federal District, as it hosts the federal government, must present even greater differentials, as long as the federal employees have the best wages among the public servants. Since *Plano Real*, there is a constant increase on these wages,

[1] *Plano Real* (freely translated as Real Plan) was a stabilization plan made in 1994 during the government of Fernando Henrique Cardoso. It was designated to target the hyperinflation and organize the public budget. The plan has accomplished its aims.

intensified on the Lula government from 2003 onwards. The decade of 2000 was marked by economic growth and greater tax revenues. These factors, under a left government of the *Partido dos Trabalhadores* (Workers Party), were decisive to increase the salaries of public servants (Gomes et al, 2012).

The information provided by the Statistical Bulletin of Employees (*Boletim Estatístico de Pessoal*) made by the Ministry of Planning shows an increase on the amount spent with public servants in all the three powers. In Brasilia (Federal District), the total expenses of the executive with its employees (active and inactive ones) has climbed from R$ 4.94 billion in 2000 to 9.25 billion in 2010 (values of 2010). On the Federal District, differently from Brazil, the inactive servants (retired ones) were more numerous than the active servants, corresponding to 58% in 2000 and 61% in 2010. This pattern suggests that the government expenditure is even more regressive and inequality creator than what was computed here, as our analysis is restricted to only the active public servants. As the Censuses do not permit the identification of the pension's recipients of the Governmental Pension System (*Regime Próprio de Previdência*), it was impossible to include this group on the study.

Also, the Government of the Federal District is the local government which better pays its employees among all Brazilian Federal Units. The Military and civil police and firemen are paid with Union money, and they receive the best wages of their categories on the whole country (MPOG, 2010). Also, the teachers in public schools have the greatest salaries in the country (Secretaria de Transparência e Controle do DF, 2013).

The household per capita income of the households with a public servant as the head has grown more than the average for Brasilia as whole. In consequence, we ought to observe a greater participation of this group in the top quantis of income. The table 2 below illustrates this phenomenon, showing an expressive increase of this group among the richest.

**Table 2 – Proportion of households headed by public servants by income quantis of total income - Federal District, 2000 and 2010**

| Quantis | Public Servants Proportion - 2000 | Public Servants Proportion - 2010 |
|---|---|---|
| **25% poorest** | 1.21% | 0.29% |
| **25 a 50%** | 5.62% | 2.71% |
| **50 a 75%** | 17.28% | 14.47% |
| **25% richest** | 23.20% | 31.56% |
| ***10% richest*** | *23.73%* | *35.91%* |
| ***5% richest*** | *23.43%* | *38.27%* |
| ***1% richest*** | *20.00%* | *38.96%* |

| | | |
|---|---|---|
| **Proportion on total population** | 11.83% | 12.29% |

Source: Censuses 2000 and 2010, IBGE. Author's calculation.

As expected, on the beginning and middle of the distribution, we can observe a drop on the proportion of public servants between 2000 and 2010. From the 25% richest onwards, there is a crescent elevation on the share of public servants, reaching its greatest levels on the top of the distribution. In 2000, 23.2% of the individuals on the top quarter lived in households whose head was a public servant. In 2010, this proportion was 31.56%. Among the top 1% in 2000, one fifth was public servants. In 2010, it reached almost two fifths, a great change for only ten years. Also, in 2000, 49% of all the public servants were among the richest quarter. It has climbed to 64.4% in 2010.

In this context, this study tries to demonstrate that the public sector generates inequalities on the Federal District. The inequality decomposition shows that inequality would have a falling trend if we have not observed the elevation on public servants' remuneration. The main determinant of income inequality, education, had an equalizer effect on the period, as long as the schooling disparities have diminished. However, inequality has slightly risen because of greater salaries on the public sector and unobserved characteristics. This unobserved characteristics can also be related to the public sector, and it will be discussed later on the results section. The subsection below analyzes the recent migration pattern on the Federal District. It will be shown that this pattern of migration tends to perpetuate inequality and is also influenced by the public sector.

### 2.2 The recent migration

Created on the end of the 1950 decade, Brasilia is still a young city for Brazilian and World standards. It was constituted by immigrants from all the regions of Brazil and even today, the Federal District is the Federation Unit with the highest proportion of habitants born in other States (IBGE, 2012). Nowadays, Brasilia is a great attractor of immigrants, qualified or not. Among the qualified ones, the public service is the main factor attracting them. According to Da Mata et al. (2007), the key characteristics for a city to attract qualified migrants are the wages, high level of education and distance to other cities. The Federal

District has all these characteristics, as it is located in the center of the country and it has the highest per capita income. Hence, by the index for net qualified immigration created by these authors using data from 2000, Brasilia is the third city in Brazil, right after São Paulo and Rio de Janeiro. Perhaps, if we apply this index to the data of 2010, Brasilia could be the first one, as long as São Paulo and Rio de Janeiro started to have net emigration on this decade (Santos and Ferreira, 2007).

Using the Census 2010, we analyze the recent migration in the Federal District. 315.244 people started to live in Brasilia between 2005 and 2010, as well as 222.632 have left it. It yields a net immigration of 92.612 people, or 3.6% of the population in 2010. However, the net immigration is positive for any quantil of income. There is a net emigration to the suburbs, in an area located outside the Federal District. This emigration is basically of poor people. The mean household per capita income of immigrants is R$ 1.869 and of emigrants is R$ 1.053. The immigration differs according to the income group. There are proportionally more immigrants among the 20% richest, and even more on the top of the distribution.

The proportion of immigrants among the poorest ones is 11.59%. Among the richest, it is 14.83%. On the top 1%, this proportion is 20.05%. The greater immigration of high income people, *ceteris paribus*, is an element which concentrates income (Santos and Ferreira, 2007).

The table 3 below shows this data:

**Table 3 – Proportion of people who was living in other Federal Unit in 2005 by income quantis – Federal District, 2010**

| | Already living on the Federal District | Living in other Federal Unit (State) |
|---|---|---|
| **20% poorest** | 88.41 | 11.59 |
| **20-40%** | 88.43 | 11.57 |
| **40-60%** | 87.01 | 12.99 |
| **60-80%** | 88.01 | 11.98 |
| **20% richest** | 85.16 | 14.83 |
| ***10% richest*** | *83.38* | *16.62* |
| ***5% richest*** | *81.92* | *18.08* |
| ***1% richest*** | *79.95* | *20.05* |
| **Total** | **87.41** | **12.59** |

Source: Census 2010, IBGE. Author's calculation.

The qualified immigration to the Federal District is highly influenced by the presence of the public administration. Brasilia hosts the three powers of the Federal government, as

well as a variety of agencies, specific authorities and state companies. As the public sector, between 2000 and 2010, has remunerated better its employees, Brasilia has become more attractive for qualified workers in search of better opportunities. The proportion of public servants among the immigrants is higher than its proportion on total population. Considering the household heads[2], 19.4% of the public servants of the Federal District in 2010 did not live there five years before. This proportion is substantially higher than the general population (12.6%). Also, among the richest immigrants, 34.5% were public servants. Finally, among all the public servants who have immigrated to Brasilia, 64.7% is among the richest fifth of the income distribution. The next section shows the methodology used to investigate the determinants of inequality on the two years. These methods can validate the great and crescent influence of the public sector on inequality in Brasilia.

[2] Constraining the sample to this group, we can access the proportion of families who have moved by different characteristics of their household heads. The public servants considered are only the stable ones, as long as it is not possible to identify the temporary ones, state companies' employees and politicians. If we could identify them, the immigration would be probably even more influenced by the public sector and its effects on inequality would be even higher.

# 3. Methodology

## 3.1. Inequality decomposition by population subgroups

The class of Generalized Entropy indexes (GE) can be perfectly decomposed using the population subgroups, differently to the Gini Index (Cowell, 2009). These indexes show the five necessary properties to measure and decompose inequality: transferability, independency, population principle, anonymity and decomposability (Litchfield, 1999). Using the GE indexes, the inequality can be divided into two parts, one with the differences between population subgroups and the other with the remaining inequality inside the subgroups. They are called *between* and *within* inequality. This method can give us the proportion of inequality coming from heterogeneities between different groups in a society (Bourguignon, 1979; Shorrocks, 1982). Total inequality, I, is equal to the sum of the two partitions:

$$I = I_{between} + I_{within}$$

This decomposition access the contribution of a specific individual characteristic to total inequality in a moment in time. Because of that, this decomposition is called static. A population subgroup is composed by people who share a common characteristic. It is possible to calculate the portion of inequality coming from the differences between genders, race/color, age group, region of the household, education, type of occupation, among others. A drawback of this method is its failure in controlling by other variables (Wan, 2004). In other words, each measure only attributes the inequality between the mean income of mutually exclusive subgroups and within the subgroups. It is not possible to use all the subgroups at once.

Formally, using the Generalized Entropy indexes, the inequality between groups is calculated as below:

$$I_{between} = \frac{1}{(\propto^2 - \propto)} \sum_{j=1}^{k} [\, f_j \left(\frac{\overline{y_j}}{\overline{y}}\right)^{\propto} - 1 \,]$$

Where $f_j$ is the proportion of subgroup j on total population. And, $\overline{y_j}$ is the mean income of the subgroup *j*. The number of subgroups is given by *k*. α is the parameter used. It is a real number which can range from -∞ to +∞, although the only parameters habitually employed are 0, 1 and 2. With α=0, we have the mean logarithmic deviation index or Theil-L. With α=1 it is the Theil-T index and with α=2, it is the half the square of the coefficient of variation (Cowell, 2009).

The result shows us the portion of inequality "explained" only by the different mean incomes of the *k* groups. The weight of each subgroup on the total is its share on the population. In the case of gender, *k* will be two and $I_{between}$ yields the portion of inequality attributed to differences of income between males and females.

The inequality within groups is defined by the equation below:

$$I_{within} = \sum_{j=1}^{k} v_j^{\propto} f_j^{1-\propto} \, E(\propto)_j$$

With $v_j$ representing the proportion of the total income that subgroup *j* holds. $GE(\propto)_j$ is the inequality measured inside each group *j=1,2,...,k*. Thus, $I_{within}$ gives us the weighted sum of inequality inside each group (Bellù e Liberati, 2006).

Cowell e Jenkins (1995) makes an intuitive meaning to the $I_{between.}$ It can be interpreted as the $R^2$ of a regression. Thus, we can find the share of total inequality explained by the differences between subgroups. Mathematically:

$$R_b = \frac{I_{\text{between}}}{I}$$

$R_b$ is the proportion of inequality which comes from subgroups heterogeneities. Because of the disadvantages of this method, we also use the regression-based decomposition. We use the most modern decomposition techniques, as the Fields' method and the Shapley value in order to have a robust estimation of each identifiable contributor to inequality. These methods are described in the next section.

### 3.2. Regression-based decomposition of inequality

The literature of regression-based decomposition has Oaxaca (1973) as a pioneer. In his work, he has measured the amount of wages disparities between men and women due to attributes and discrimination. This approach was an innovation for the time, but it focused in calculating the percentage of the wage difference related to this two distinguished explanations. In this paper, we are more interested in using inequality indexes and disaggregate the effect of each variable. The literature has been slightly developed until the nineties. In that decade, new papers using quantile regression and non-parametric statistic were published, nevertheless the results has led to weak conclusions (Wan, 2002).

According to Wan (2002) and Wan and Zhou (2004), in the end of the nineties, the interest of the researchers had a turn over into the direction of quantifying the contribution of many determinants to total inequality, instead of analyzing the differences between groups (males vs. females, whites vs. blacks…). The papers of Fields (1998), Fields and Yoo (2000), Shorrocks (1999) and Morduch and Sicular (2002) were the basis for the modern regression-based decomposition. Wan (2002; 2004) has given important methodological and empirical contributions.

The starting point for any regression-based decomposition is to define an income generating function. In this case, income will be the dependent variable, in linear or logarithmic form. The independent variables are the determinants of income. These variables are related to individual characteristics, such as education, experience/age and also dummies of gender, race, type of occupation, household region… We can include n variables we think can affect the household per capita income, and, consequently, its distribution (Wan, 2002).

Each variable, thus, has a contribution to the total income. On the regression analysis, this contribution is used as an income factor (Morduch and Sicular, 2002). The total income is the sum of the income generated by all the factors and the residual term. On the next subsection, we explore the Fields' method, which is used by Fields and Yoo (2000), Litchfield (2001) and Salardi (2005), the last two for Brazilian data.

### *3.2.1 The Fields' Method*

Fields (1998) has developed the method which consists of using a regression of the logarithmic of income. After the estimation, we measure the covariance between each variable and the log of the income. The method was applied by Fields and Yoo (2000) to South Korea's labor market in order to access the percentage explained by each variable in the inequality of wages. The estimated regression has this functional form:

$$\ln(Y) = \beta_0 + \sum_k \beta_k x_k + \varepsilon = \sum_j a_j Z_j$$

with $a_j = [\beta_0\ \beta_1\ \beta_2 \ldots \beta_j\ 1]$

and $Z_j = [1\ x_1\ x_2 \ldots x_j\ \varepsilon]$

The author proves that the inequality of the logarithmic of income can be perfectly decomposed using its variance. He argues that this result would be the same with any inequality index. The contribution of each variable is simply its covariance with the log of income normalized by the total variance, such as follow:

$$Sj\ (ln\ Y) = \frac{cov\ [aj\ Zj, \ln Y]}{\sigma^2(\ln Y)} = \frac{a_j\ *\ \sigma(Z_j)\ *\ cor\ [Z_j \ln Y]}{\sigma(\ln Y)}$$

Then,

$$\sum_j Sj\ (ln\ Y) = 100\%$$

The total explained is:

$$\sum_j^{J-1} \frac{cov\ [aj\ Zj, \ln Y]}{\sigma^2(\ln Y)} = R^2\ (\ln Y)$$

And the proportion of each variable on the total is:

$$p_j\ (\ln Y)\ = \frac{Sj\ (\ln Y)}{\sum_j Sj\ (\ln Y)}$$

Nevertheless, the method has the flaw of decomposing an improper measure of inequality, because the variance does not respect the independency principle. The author argues that the covariance between an explanatory and explained variable also changes with the scale, which makes possible the comparability between two moments of time. Although, Morduch and Sicular (2002) and Wan (2002; 2004) criticize this approach, arguing that this decomposition is problematic, as it uses the log of income instead of the real income. The analysis of the log of income keeps the ranking of incomes unchanged, but it makes the highest incomes not to have the right weight on the distribution. In any case, we use the Fields' method together with the most modern decomposition technique, the Shapley value. Every method leads to the same conclusion, as shown in the results.

Morduch and Sicular (2002) use a linear equation in order to decompose inequality. Their results are very sensible to the index employed. In addition, the residual of the estimations is quite large (more than 50% in almost all cases). Because of that, it is not advisable to use linear equations to reach policy conclusions. Facing these pitfalls, Wan (2002; 2004) refines the method using the Shapley value, earlier integrated to the inequality analysis by Shorrocks (1999). On the next subsection we present these improvements and also the applicability to data.

### *3.2.2 The Shapley value*

The great contributions of the works of Wan (2002; 2004) are their capacity of generalizing the previous methods for any functional form of income generating function, as well as inequality index. Furthermore, it treats some inconsistencies neglected before. The greatest problem identified by Wan is related to the residual. As it has zero mean, it does not affect the mean of the dependent variable, but it affects its density and, in consequence, its distribution. Another problem is related to the constant. By the principle of transferability of Pigou-Dalton, if the constant is positive, the measured inequality would be underestimated,

because the income of everyone would have been increased by an equal amount. With the purpose of target these questions, Wan uses the following equation:

$$Y = F(X) = \beta_0 + Y(X) + \varepsilon = Y^* + \varepsilon$$

On this equation, Y can be in the linear or logarithmic form, but the last one is preferable. Y(X) corresponds to the estimated coefficients for each independent variable. Y* is the entire portion explained by the model ($\beta_0$ + Y(X)) and ε is the error term.

In order to divide the explained proportion and the contribution of the error term, we must treat the latter as a residual of the former (Wan, 2002). The question to be answered is: if the error term did not exist, what would be the inequality measured? Applying an inequality operator (any inequality index), I, we have:

$$I(Y|\varepsilon=0) = I(Y^*)$$

Therefore, the explained proportion is the inequality of the predicted values of the regression, which is necessarily lower than the inequality of the observed values. The contribution of the error term is simply the difference between these two measures. Thus, the contribution of ε ($C_E$) to total inequality is given by:

$$C_E = I(Y) - I(Y^*)$$

In the case of the constant, its contribution is also defined as a residual. We need to measure the inequality as if the constant was zero. Following the equations above, it gives us:

$$I(Y^*|\ \beta_0=0) = I[Y(X)]$$

And the contribution of the constant is:

$$C_{\beta 0} = I(Y^*) - I[Y(X)]$$

Hence, we can split the total inequality into three parts:

$$I(Y) = C_{\beta 0} + I[Y(X)] + C_E$$

The percentage of contribution of each term is simply its value divided by the total inequality - I(Y). In the case of a log-linear equation, the contribution of the constant is zero. It happens because the decomposition is made after calculating the exponential of the estimated logarithmic income. Disaggregating F(X) and applying the properties of exponential, we have:

$$F(X) = EXP(\beta_0 + \beta_1 X_1 + ... + \beta_k X_k + \varepsilon) = EXP(\beta_0) * EXP(\beta_1 X_1) * ... * EXP(\beta_k X_k) * EXP(\varepsilon)$$

Thus, the constant is only a scalar multiplying all the variables of the equation. By the principle of scale independency, we know that the inequality is not affected, and, under this case, the contribution of the constant is null. This paper uses this functional form.

In order to access the explained proportion of each independent variable, we use the Shapley value, as described by Shorrocks (1999). In this method, it is applied the notion of cooperatives games to the analyses of inequality. In a seminar paper of the winner of 2012's Nobel Prize, Shapley (1953) argues that in a cooperative game with n individuals, the payoffs will differ from player to player, because each one has a different bargaining power (coalitions) when playing the game. Analogously, concerning inequality, each determinant of income has a particular importance for the total, and it also relates differently with each other.

The Shapley value gives us the marginal impact of each independent variable. We define inequality as a function of s variables: k explanatory variables and the error term (s=k+1). Formally we have: $I=(X_1, ..., X_j, ..., X_k, \varepsilon)$ or $I = (X_1, ..., X_j, ..., X_k, X_s)$, and I is an inequality index. The contribution of each variable is given by the difference between total inequality and the inequality measured when this variable is 'excluded' (replaced by its mean – the inequality of this variable is null)[3]. The structure of the model is (K,F). K is the set of factors and F is the function used to determine the income.

Naming as S the set of s variables, I[F(S)] is the inequality when the j-th variable is equally distributed. I[F(1)] is the inequality when one variable is equally distributed, I[F(2)] measures it when two variables are equally distributed. I[F(K)] is the total explained

[3] There are two distinct methods: replace the variable by its mean or by zero. The first eliminates the inequality generated by the variable and the other removes the variable. Removing the variable is quite problematic, because the results are more volatile and less intuitive (Sastre e Trannoy, 2002). Thus, we opted for the first method. It provides basically the value of inequality if the j-th determinant was equally distributed, or, in other words, it did not have any impact on the income distribution.

inequality (I without any explanatory variable, only ε). I[F(Ø)]=0, it means the inequality is zero if all determinants were replaced by theirs means.

However, the marginal contribution differs as long as we change the order of replacement. For example, if $X_j$ was the first to be replaced by its mean, its contribution will be distinct from when it was the seventh. Thus, we must compute the marginal contribution in every possible order and calculate the mean for each variable. We define $M = \{1, ..., m, ..., s\}$ as the set with the sequence of replacements made. We must compute the effect of only one replacement (one variable at once) and many. We must be aware that the position of replacement of each variable affects all the other results, including the error term. The maximum number of possible permutations is s!. Hence, we define $C_j$ (K, F) as the marginal contribution of the j-th determinant of inequality. We have:

$$C_j\ (\mathrm{K},\mathrm{F}) \ = \ \sum_{m=0}^{s} \sum_{j \in \mathrm{S}} \frac{(s-m)!\,m!}{s!} \ [\ I\big(F(S)\big) - I(F(S\ - \{j\}))\ ]$$

The last term yields the inequality measured if we replace the factor j from the set of factors S. Thus, the marginal contribution of each variable is equal to the mean of the *m* replacements made. This is the Shapley value, which yields an exact and independent decomposition (Shorrocks, 1999). The number of subsets composed by the variables is $2^s$. If we have three variables, $x_1$, $x_2$ and $x_3$, we must compute the inequality replacing each of these variables by its mean and all the possible interactions between them. Thus, we have $2^3$ computations: I[F($x_1$)], I[F($x_2$)], I[F($x_3$)], I[F($x_1 \cup x_2$)], I[F($x_1 \cup x_3$)], I[F($x_2 \cup x_3$)], I[F($x_1 \cup x_2 \cup x_3$)] and I[F(Ø)]. The impact of a variable is yielded by the mean of its marginal impact in each order of replacement (Charpentier and Mussard, 2010). When we replace $x_1$ on the first position, the marginal impact on this position is itself. When we replace it on the second position, its impact is the mean of the impacts when $x_2$ is first replaced and when $x_3$ is first replaced. On the third and last position, the impact of $x_1$ is also itself. Thus, the contribution of $x_1$ is the mean of the marginal impacts in every position.

Shorrocks (1999) and Sastre and Trannoy (2002) prove that the Shapley value is a symmetric decomposition. It is also sensible to the inequality index used. However, Wan (2004) and Baye and Epo (2011), employing this method to Chinese and Cameroons data with a log-linear function, found that the sensibility to the index is not high, as it cannot change the results qualitatively. In the opposite, the linear functions used by Morduch and Sicular (2002)

have this problem. Araar and Duclos (2007) have developed a package for Stata ® which computes the Shapley value. It demands great computational capacity, as each variable has to have its marginal impacts computed. This paper uses the implementation made by these authors, as it is available for free. The next section present the results, firstly of the static decomposition by population subgroups and then of the regression-based decomposition.

# 4. Results

## 4.1. Inequality decomposition by population subgroups - Brasilia, 2000 and 2010

In order to understand the major contributors to the inequality in Brasilia (Federal District) in 2000 and 2010 we start decomposing the inequality of household per capita income by the Generalized Entropy indexes. It also shows us guidelines about the reason why the inequality in Brasilia has kept high while Brazil as whole has experienced a strong reduction in these indicators.

The argument that the main factor which kept high levels of inequality was the raising income of public servants is corroborated by this methodology. Also, the results of the next subsections, using the regression-based decomposition, are even stronger and conclusive.

We opted to use the characteristics of the household head to divide the population. Litchfield (2001), Salardi (2005) and Ferreira et al. (2006) used this approach. As long as we are measuring the inequality of the household per capita income - all the individuals in the Family has the same income - the smallest unity of analyses is a person who lives in a household. The household head is understood as having representative characteristics. Hence, the results must be interpreted as the percentage of inequality coming from differences between household's heads[4]. We used six attributes to divide the population into subgroups. They are:

*1) Education of household head*

Perhaps this characteristic is the most important one to explain the income inequality in any place of the World (Bourguignon and Morrison, 2002; Becker, 2005). It is known that the income is higher according to the level of education of a person, on average. Also, we would expect families with an educated head to be composed by a more educated partner as well as by an offspring in the process of education. We believe that the level of education explains a considerable part of total inequality, especially in Brazil, as long as there are still few people with graduation and post-graduation studies comparing to OECD countries (Barros et al., 2007). In order to create the subgroups of this variable, we follow Ferreira et al.

[4] Each household head has a weight equal to the size of his family multiplied by the weight of the household, which was attributed by IBGE in the survey.

(2006), but we add the last interval. There are 6 subgroups: illiterate, from 1 to 4 years of education (first years of Primary Education), from 5 to 8 years (last years of Primary), from 9 to 11 years (High School), from 12 to 15 years (Undergraduate studies) and 15 or more (Post-graduation studies).

*2) Age of household head*

The decomposition by this characteristic can access the inequality of age subgroups. Consequently, we can see if there are disparities between households whose head is young or old. We also divided this variable into 6 subgroups: below 25 years old, between 25 and 34, 35 and 44, 45 and 54, 55 and 64 and 65 or more.

*3) Race/color of household head*

Here, the inequality is decomposed by differences of household per capita income according to the race/color of the household head. Brazil, because of its past of slavery, still presents racial inequality, and it also persists in the big cities (Garcia, 2006). The sample was separated into three subgroups of race/color: whites, blacks and mixed, and others. The last one includes indigenous, Asiatics and non-responses. This subgroup represented only 1.2% of the population in 2000 and 1.8% in 2010.

*4) Gender of household head*

The household head can be a man or a woman. We would expect the household per capita income to differ, given the vulnerability of families constituted by single mothers and the lower wages of women in the labor market, a worldwide phenomenon (Gustafsoon e Johansson, 1999).

*5) Family size*

We divided the households into six groups, according to the number of residents. It is divided into households of one person, two, three, four, five and six or more. The household per capita income strictly decreases with family size, making the disposable income for each habitant diminish. It is a result of more children and old people living in the same household, as they do not work. In addition, poorer families present more habitants on average, given their budget constraint (Ferreira et al., 2006). Because of that, differences between those groups must have a non-negligible impact to total inequality.

*6) Type of occupation of household head*

With this variable, we can analyze if there are inequality between households headed by people with different occupations. The subgroups were constructed considering the employer and the condition of the job (formal or informal). This variable access the contribution to total inequality accrued from income differentials between public servants[5] and other groups. We have divided the sample into nine subgroups: civil or military public servant, formal worker in the private sector, informal worker in the private sector, formal domestic worker, informal domestic worker, autonomous worker, employer, other (non-remunerated jobs) and unoccupied (unemployed or retired)[6].

*7) Region of household*

Finally, we divided the population into subgroups according to the region of the household. The Federal District of Brazil was divided into 19 Administrative Regions and its rural zone, following the Census of 2000 (details on the appendix A). The urban segregation is still very present on the Brazilian metropolises. The way the space is occupied is considerably influenced by the income (Garcia, 2006). For that reason, the decomposition will give us how much the local differences contribute to inequality in Brasilia.

Below, on Table 4, we show the inequality of 2000 and 2010 measured by the Gini index and by three index of the Generalized Entropy class. An explanation must be done: on the Census 2010, because of a mistake of the staff, some non-declared incomes were reported as zero. 4% of the sample had null income on this year. In order to identify these 'false poor', we investigated whether the household had relevant goods, such as fridge, automobile, computer, washing machine and television. If the household with null income had two or more of those items, it would have its income changed to missing and would be excluded from the sample. After that, still 2.8% of the sample had null income. On this paper, we opt to use the sample with 2.8% of null incomes, as long as all those households were headed by an unoccupied person or informal worker. This result makes us to believe that their income is truly zero or close to zero, given the vulnerability of those groups. It is reasonable that the

[5] The public servants analyzed are only the ones working directly to the government under the special regime. The employees of state companies and temporary public servants are not included. The questionnaires of the Census do not specify these two categories. On the sample, they are included in the group of formal workers of the private sector.

[6] The household head who works in a non-remunerated job or is unoccupied can earn income from other sources such as pensions, interests or transferences. Furthermore, as we measured the household per capita income, the earnings of the other members of the family are taken into account in order to compute the mean income of the household.

Federal District have people living in extreme poverty, with per capita income close to zero (which, for practical matters, does not have an important impact when reported as zero). All the indexes have shown an increase on the Federal District's inequality [7]. We can affirm that the entropy of the income has risen[8]. The decomposition analyses indicate possible causes for this.

**Table 4 –Gini Index and Indexes of the Generalized Entropy Class for the Federal District of Brazil, 2000 e 2010**

| | 2000 | 2010 | Total Variation | Percentage Variation |
|---|---|---|---|---|
| **Gini** | 0.627 | 0.637 | 0.010 | 1.6% |
| **GE(0)** | 0.771 | 0.846 | 0.075 | 9.7% |
| **GE(1)** | 0.789 | 0.850 | 0.061 | 7.7% |
| **GE(2)** | 2.210 | 3.530 | 1.32 | 59.7% |

Source: Censuses 2000 and 2010, IBGE. Author's calculation.

On Table 5, we can see the percentage estimated which each attribute contributes for the inequality in 2000 and 2010, based on the three most important indexes of the Generalized Entropy Class. The GE(0) and GE(1) show close results, whereas GE(2), by its mathematical construction, has a tendency of presenting a lower $I_{between}$. This is more common in unequal societies (Litchfield, 2001). As long as the differences on the top of the distribution have risen, the inequality measured by the GE(2) has climbed, as shown. Also, as the differences between the groups have not changed so strongly on these ten years, the percentage explained of the GE(2) had dropped more than the others indexes. Because of that, the results of the GE(0) and the GE(1) are more robust.

[7] The Theil-L and the Theil-T do not compute null incomes. In order to target this pitfall, we followed the procedure used by Eble (2007). We have imputed an income equivalent to 1% of the mean income to every person with zero income. The Census 2000 does not have people with null incomes, because IBGE has published the data after making imputations.

[8] If we calculate the indexes of 2010 using the sample without any null income, we would have Gini=0.628, GE(0)=0.759, GE(1)=0.823 and GE(2)=3.419. Only GE(0)would have reduced in relation to 2000, simply because it gives a stronger weight to the lower bound of the distribution. These estimations can be considered as the lowest level of inequality for the Brazilian Federal District in 2010. It is still superior (or at least not inferior) to the values of 2000.

**Table 5 – Static Decomposition by Population Subgroups - Federal District of Brazil, 2000 e 2010**

| Subgroup | Explained Proportion 2000 | | | Explained Proportion 2010 | | |
|---|---|---|---|---|---|---|
| | GE(0) | GE(1) | GE(2) | GE(0) | GE(1) | GE(2) |
| **Education** | 41.9% | 43.9% | 20.2% | 36.4% | 37.8% | 11.1% |
| **Age** | 5.3% | 4.8% | 1.6% | 4.3% | 4.1% | 1.0% |
| **Race/color** | 9.7% | 9.1% | 3.2% | 8.2% | 8.2% | 2.0% |
| **Gender** | 0.7% | 0.7% | 0.2% | 1.3% | 1.2% | 0.3% |
| **Family size** | 7.4% | 8.1% | 3.4% | 11.2% | 11.2% | 2.9% |
| **Type of occupation** | 12.9% | 13.6% | 6.3% | 14.2% | 15.4% | 4.5% |
| **Region** | 40.1% | 41.5% | 18.9% | 32.5% | 35.1% | 10.7% |

Source: Censuses 2000 and 2010, IBGE. Author's calculation.

As expected, the variables which better describe the nature of the inequality in Brasilia are region of the household and education of the household head. According to the indexes GE(0) and GE(1), these two variables corresponded, each, for more than 40% of the total inequality in 2000. This results follow Ferreira et al. (2006) and Azzoni and Neto (2011), which show a trend of smaller regional disparities and less accentuated returns to education in Brazil, although still high. We could say that the Federal District confirm this for the case of education. Considering the regional differences between 2000 and 2010, new neighborhoods were created. This pattern does not allow us to conclude anything, because the regions used were the same of 2000. In consequence, the internal inequality of these regions has risen, indicating more disparities. In order to compare the regions, we allocated the new neighborhoods into the areas of the 2000's jurisdiction. If we had disaggregated the new localities, the regional differences would certainly be greater.

The age of the household head had a reasonable relevance on total inequality. It has shown a slightly reduction between 2000 and 2010. This demonstrates that the intergenerational income differences are not so high for individuals who already constituted a family, but still has a contribution to inequality. Income is strictly crescent considering the age of household head. The youngest ones (below 25 years old) show an income lower than 50% of the average. On the other hand, the oldest ones (more than 65) have an income 40% higher than the average. This pattern has not had major changes between 2000 and 2010.

The race has been important to explain a minor part of the inequality in both years, but also showing a lower percentage in 2010. This result, as long as it does not control for other variables is super-estimated. It will be clear on the next section, with the regression-based decomposition. There was an expressive increase on the proportion of household heads who declared themselves as blacks or mixed. In 2000, 51% of the population was living in a household whose head was black or mixed, and it has risen to 57.7% in 2010. Whites were 47.5% in 2000, falling to 40% in 2010. This is an indicator of a greater perception of the racial issue, which was discussed by public campaigns against racism and also by the policy of quotas in the University of Brasilia. Nevertheless, the income of the black and mixed people is still considerably lower, only 45.5% of the whites in 2000 and 46.9% in 2010, less than half.

Comparing the two years, we can noticed that the majority of the variables has shown a weaker capacity to explain the total inequality in 2010, with the exception of gender, family size and type of occupation. The gender of the household head was not so relevant for the inequality and it accounts to less than 1.5% in any of the years. The slightly increase in the percentage attained to this variable is probably a consequence of more women declaring themselves as household's head. They were only 32.5% in 2000, but 43.1% in 2010. This phenomenon was already noticed by Salardi (2005) for Brazil as a whole. It indicates a better recognition of the female figure as the person in charge for her family, even if she has a monetary income lower than her partner.

The family size seems to be more relevant to the total inequality in 2010. However, the regression-based decomposition, as it controls for other elements, does not show a change in this variable between the two years. The average family size was 3.71 people in 2000 and 3.3 in 2010. If we divide by income quantil, among the 20% richest, the mean number of habitants has dropped from 3.14 to 2.69. Among the 20% poorest, it has dropped from 4.61 to 3.96. On percentage, the drop was basically the same (14.3% for the rich and 14.1% for the poor). This result does not allow us to conclude that a change on family size has had a relevant contribution to the maintenance of a high inequality level in the Federal District.

Lastly, the type of occupation of the household head was important to explain part of the inequality in 2000 and 2010. The percentage measured by the GE(0) and GE(1) has raised from 12.9% and 13.6% to 14.2% and 15.4%, respectively. The percentage measured by the GE(2) has shrank, because of the reasons already mentioned. If we disaggregate for each subgroup of occupation, we can see that the most important one for the inequality in 2010 was the public servants'. The GE(0) (Theil-L) and the GE(1) (Theil-T) are constructed in way that

the subgroups with income above the mean will contribute positively to the index, while the poorer subgroups will contribute with a negative amount. Considering type of occupation, only public servants and employers have an income above the average. Thus, looking to the contribution of the public servants' subgroup, we see that it rose from 7.7% to 11.6% of the Theil-L between 2000 and 2010. Analyzing the Theil-T, the increase was even bigger, from 12.7% in 2000 to 26% in 2010. All the other subgroups of any variable have had a similar or reduced impact between 2000 and 2010.

The inequality between households with heads in different occupations has increased in those ten years. The changes in the income of the poorer subgroups were small. On the other hand, among the richest - employers and public servants, it has had important variations. The former has had their income proportionally reduced when compared to the population as a whole (from 3.28 times the average to 2.90). The latter, on opposite, had showed an elevation (from 1.67 times the average to 2.24). Also, the proportion of employers had dropped, from 3.4% to 1.9%. The inequality coming from the types of occupation was, in 2000, greatly related to the group of employers. In 2010, the group of the public servants has had the biggest weight to inequality by type of occupation, since it has showed a considerable rise on incomes and a slightly one on the population proportion, from 11.8% to 12.3%.

Nevertheless, the static decomposition is insufficient to affirm that the rise in the remuneration of the public sector was the main explanation for the high level of inequality in Brasilia. As said, this analysis does not control for other variables. The regression-based decomposition shows more conclusive and robust results which confirm the ones presented here. The next section estimates a regression with the determinants of income and the subsequent one shows its decomposition.

### 4.2 Estimating the determinants of household per capita income

On this section, we estimate the determinants of the household per capita income. We make a log-linear regression which is the basis to the decomposition presented on the next section. The functional form is basically an income generating function following Mincer (1974). This regression gives us the returns of years of school, age/experience and other individual characteristics on income. The dependent variable is the natural logarithmic of the household per capita income. We did not need to separate the households into subgroups, as in

the last section. Instead, we used continuous, discrete and dummies variables in its original form. This type of regression, as it is log-linear, reports the marginal impact of each independent variable in percentage, approximately (Wooldridge, 2002). Also, it is better adjusted to describe income and also allow more robust and less volatile decompositions (Wan, 2004).

Providing the decomposition by the Shapley value demands a great computational capacity, it is not possible to include all the dummies of region. In order to avoid loss of information, we opted to aggregate the regions into three big groups according to their level of income. They are: the rich ones, the middle income ones and the poor ones[9]. The group for comparability is the one with the poor regions.

In order to analyze the household per capita income, the estimated equation uses the characteristics of the household head. As said, the income can vary by differences on education, age, gender, race and others. These variables can also be correlated to the fact of being a public servant or not. The regression yields the increment on income of only one of those characteristics, *ceteris paribus*.

On the equation, instead of nine subgroups of type of occupation, we used only four subgroups. It was made because of the necessity of estimating a shorter regression in order to decompose it later. The four subgroups are: 1) Civil or military public servants, 2) Private sector workers, 3) Employers and 4) Unoccupied and informal or domestic workers[10]. The last one is the base group. We opted to put the unoccupied and the informal and domestic workers together because they presented similar mean income on both years, which is smaller to the total average. The estimated model is:

$$Ln\,(Y_{pc}) = \beta_0 + \beta_1\,Education + \beta_2\,Age + \beta_3\,White + \beta_4\,Woman + \beta_5\,Family\ Size + \beta_6\,Public\ Servant + \beta_7\,Private\ Sector\ Worker + \beta_8\,Employer + \beta_9\,Rich\ Region + \beta_{10}\,RA\ Middle\ Income\ Region + \varepsilon$$

*Education* refers to the years of schooling completed by the household head. *Age* is his age in years[11]. *White* is a binary variable with value 1 for whites and 0 otherwise. *Woman*

[9] For each year, the regions with mean income below 60% of the average of the Federal District were classified as poor. The regions with income between 60% and 150% were classified as middle income. Lastly, the regions with mean income above 150% of the average were considered rich. The regression results with all the regional dummies are very similar to the ones here (coefficients equal even on the third decimal).

[10] The regression results with all dummies of type of occupation do not differ qualitatively from the results here.

[11] We opted not to include this variable in the quadratic form, as we are using the household per capita income and it is strictly crescent to the age of the household head in the years 2000 and 2010.

is a dummy for females. *Family size* is the number of people living in the same household (on the analysis, they have the same income). *Public Servant*, *Private Sector Worker* and *Employer* are dummies with unitary value for each of these types of occupation. Lastly, there are two dummies for the region of the household, the first for rich ones and the second for middle income ones.

The Table 6 below corroborates with mostly of the results of last section. The coefficients of *Age, White* and *Region* have diminished. In 2000, the increase on income was of 2.1% for each year old, while this value was 1.8% in 2010. The difference between the income of whites and non-white has dropped from 15.8% to 14.8%. The regional differences fell relatively. The habitants of the rich regions had income 90% greater than the poor ones in 2000 and 84% in 2010. Concerning women, we have observed a slightly reduction on the disparity with men. The effect of *Education* has kept statistically the same, but the latter decomposition shows a less accentuated impact of this variable on inequality.

**Table 6 – Determinants of household per capita income by characteristics of the household head. Brasilia - Federal District of Brazil, 2000 and 2010.**
**Dependent Variable: Logarithmic of the household per capita income**

| | 2000 | 2010 |
|---|---|---|
| **Intercept** | 4.2515 | 5.3193 |
| **Education** | 0.0976 | 0.1007 |
| **Age** | 0.0213 | 0.0182 |
| **White** | 0.1578 | 0.1475 |
| **Woman** | -0.1798 | -0.1694 |
| **Family Size** | -0.1612 | -0.1647 |
| **Public Servant** | 0.4984 | 0.6150 |
| **Private Sector Worker** | 0.2098 | 0.1234 |
| **Employer** | 0.9027 | 0.7405 |
| **Rich Region** | 0.9043 | 0.8381 |
| **Middle Income Region** | 0.3897 | 0.3303 |
| | | |
| **F Test** | 7,775 | 4,701 |
| **$R^2$** | 0.599 | 0.585 |
| **Observations** | 52,634 | 33,832 |

Note 1: All the estimations were significant at the 1% level
Note 2: We have used the method of Weighted Least Squares, on which each observation has a weight attributed by IBGE multiplied by the number of habitants of the household.

*Family size* has continued with the same coefficient on both years, around 16% of reduction for each additional habitant. Analyzing the types of occupation, the per capita income of the households headed by public servants has grown considerably, from a 'premium' of around 50% in 2000 to 61.5% in 2010. Also, the 'premium' for the employers has fallen between 2000 and 2010, which is in accordance to the previous results.

Completing the analysis, the next subsection shows the decomposition of inequality using the regression above.

### 4.3 Regression-based decomposition of inequality

On the regression-based decomposition, we use the Fields method and the Shapley value with the Gini index, the GE(0) (Theil-L) and the GE(1) (Theil-T). Wan (2004) and Wan and Zhou (2004) have made the decomposition using the Shapley value by these three indexes and verified that the sensitivity is not high. They concluded that the results are robust to the index used. Our analyses reach the same conclusion, since the results do not change qualitatively.

We decompose the household per capita income based on the previous regressions. The results yield the proportion of inequality attributed to each variable. This method shows us the relative contribution of each variable controlling by all the others. With this decomposition, we can see the changes on the inequality determinants on the two years. We can access the impact on inequality coming from the higher income of public servants. A positive effect would indicate the better remuneration in the public sector is inequality-increasing.

The Table 7 below shows the contribution of each variable to total inequality by the Fields' method and the Shapley value using the Gini index, the Theil-L and the Theil-T. Education has had the greatest impact on inequality. However, all estimations have indicated a reduction on the share of education on inequality between 2000 and 2010. It was about 26% in 2000, falling to 22% in 2010. This result endorses the decomposition by population subgroups, indicating that there is less inequality coming from differences on education attainment. It demonstrates the great impact of universal education on long run inequality. Also, it has important implications to public policies: providing education of all levels to as

many people as possible not only elevates the human capital of a country but also make it less unequal. Studies about the returns to education in Brazil show that the widespread of Primary Education during the nineties and the greater number of people with higher education are important factors to explain the smaller returns of this variable and the reduction on inequality on the posterior years (Azzoni and Neto, 2011; Ferreira et al., 2006).

**Table 7 – Regression-based decomposition of household per capita income by characteristics of the household head: Brasilia – Federal District of Brazil, 2000 and 2010**

| | Fields | Shapley Value | | | | | |
|---|---|---|---|---|---|---|---|
| | % | Gini | % | GE(0) | % | GE(1) | % |
| **2000** | | | | | | | |
| Education | 27.00 | 0.1611 | 25.69 | 0.2052 | 26.59 | 0.1981 | 25.08 |
| Age | 4.77 | 0.0602 | 9.59 | 0.0471 | 6.11 | 0.0502 | 6.36 |
| Race/color | 2.27 | 0.0173 | 2.76 | 0.0183 | 2.38 | 0.0175 | 2.22 |
| Gender | 0.33 | 0.0060 | 0.95 | 0.0032 | 0.42 | 0.0037 | 0.46 |
| Family Size | 7.41 | 0.0415 | 6.62 | 0.0429 | 5.56 | 0.0376 | 4.77 |
| Type of Occupation | 6.30 | 0.0633 | 10.09 | 0.0476 | 6.18 | 0.0515 | 6.52 |
| *Public Servant* | *3.66* | *0.0207* | *3.31* | *0.0206* | *2.67* | *0.0104* | *1.32* |
| *Private Sector Worker* | *0.12* | *0.0044* | *0.70* | *-0.0035* | *-0.45* | *-0.0080* | *-1.01* |
| *Employers* | *2.37* | *0.0241* | *3.84* | *0.0305* | *3.96* | *0.0491* | *6.21* |
| Region of the household | 10.24 | 0.0635 | 10.12 | 0.0805 | 10.43 | 0.0758 | 9.59 |
| Residual | 41.83 | 0.2285 | 36.42 | 0.3268 | 42.35 | 0.3553 | 45.00 |
| **Total** | **100** | **0.628** | **100** | **0.772** | **100** | **0.790** | **100** |
| **2010** | | | | | | | |
| Education | 25.83 | 0.1447 | 22.70 | 0.1881 | 22.23 | 0.1854 | 21.82 |
| Age | 4.43 | 0.0600 | 9.41 | 0.0471 | 5.57 | 0.0437 | 5.14 |
| Race/color | 2.07 | 0.0136 | 2.13 | 0.0144 | 1.70 | 0.0159 | 1.87 |
| Gender | 0.70 | 0.0060 | 0.94 | 0.0039 | 0.47 | 0.0041 | 0.49 |
| Family Size | 7.60 | 0.0378 | 5.93 | 0.0401 | 4.85 | 0.0458 | 5.39 |
| Type of Occupation | 7.72 | 0.0752 | 11.80 | 0.0697 | 8.24 | 0.0614 | 7.23 |
| *Public Servant* | *6.93* | *0.0546* | *8.57* | *0.0627* | *7.42* | *0.0508* | *5.98* |
| *Private Sector Worker* | *-0.22* | *0.0075* | *1.18* | *-0.0084* | *-0.99* | *-0.0151* | *-1.78* |
| *Employers* | *1.01* | *0.0131* | *2.05* | *0.0154* | *1.82* | *0.0257* | *3.03* |
| Region of the household | 8.33 | 0.0448 | 7.04 | 0.0564 | 6.66 | 0.0581 | 6.84 |
| Residual | 43.31 | 0.2552 | 40.05 | 0.4253 | 50.28 | 0.4354 | 51.24 |
| **Total** | **100** | **0.637** | **100** | **0.846** | **100** | **0.850** | **100** |

Source: Censuses 2000 and 2010, IBGE. Author's calculation.

The differences between races/colors and genders have also shrank from 2000 to 2010. We can observe a reduction on the contribution of these variables to total inequality. It means that if whites or non-whites would have the same income, the inequality would fall less in 2010 than in 2000, which indicates smaller differences between the groups in 2010[12]. The age of the household head seems to have preserved its contribution, varying between 4.5% and 9.5% according to the method employed. We could say the distribution of income have not changed relatively to age. The family size seems to have diminished its contribution, nevertheless the ambiguous result of the Fields' method. The regional differences, as in the decomposition by population subgroups, have been reduced. The proportion of inequality explained by this characteristic dropped from 10% in 2000 to 7% in 2010. But, as explained before, this result is not straightforward, given the new geographic configuration of Brasilia in 2010. The residual increased from 2000 to 2010. Perhaps, part of this elevation has occurred because we used the same regions of 2000 in 2010.

Undeniably, the inequality coming from differences on the type of occupation has climbed on the period. Using the decomposition of the Gini index by the Shapley value, it has moved from 10.1% to 11.8%. This variation was almost completely due to the public sector remuneration. In 2000, it had an impact of 3.31% to total inequality, achieving 8.57% in 2010. Using any method, the contribution of the group of public servants to inequality has climbed. Making a hypothetic exercise, if the absolute contribution of this group to the Gini index has kept the same, the inequality in 2010 would not be 0.637, but 0.603, 5.3% less.

The contribution of the employers, the other group with income above the average, has fallen between 2000 and 2010 on all estimations. The elevation on the income of public servants was the main factor to keep (and also slightly elevate) the inequality of the household per capita income on the Federal District of Brazil between 2000 and 2010. The greater remunerations of these employees by the federal and regional public administrations cause

[12] The small results on both years can be triggering in a first moment, as long as there are large differences on the mean income of whites and non-whites or women and men. However, the correct interpretation of the method can enlighten this topic. First, it is important to notice that the inequality between the households headed by women is slightly greater than the inequality on the households headed by men (Gini index of 0.637 against 0.632 in 2010 – full sample: 0.637). In 2010, the inequality within non-whites was 0.605 and between whites was 0.621. There are great disparities inside each group. In order to compute the contribution to total inequality using the Shapley value, we replace the values of the variable by its mean. Thus, we measure the inequality as if men and women or whites and non-whites had the same mean income. Or, in practice, we add a value to every observation of the poorest group, making this group to have the same mean income of the richest group. After it, we measure the new value of inequality. As long as the differences inside the groups are already high, the inequality would not change much. For the case of the dummies of type of occupation, the distribution of income within each group is considerably different. It makes the impact of those groups on inequality higher. For example, the Gini within the group of public servants is only 0.481 in 2010, showing that their income is more equal and concentrated in the upper tail of the distribution.

and perpetuate the inequality.

The residual term is greater on the year 2010. This is probably a reflex of the greater inequality on the income of other sources, which has climbed from 0.654 to 0.727, measured by the Gini index. This elevation has taken part in a decade of more programs of direct cash transfer. Hence, it must have had greater concentration of properties and/or deterioration on the distribution of pensions (the latter is highly influenced by the public sector). Certainly, the income from other sources shows some correlation with observable characteristics of the individual, which was accessed by the decomposition. Nevertheless, the income from other sources seems to be less correlated to identifiable individual characteristics, which has elevated the residual term by any method used.

On the Appendix B, we apply this method to the formal labor market of Brasilia. The results support the ones here described. They show greater disparities on wages, in part caused by the public sector.

## 5. Conclusion

As seen, Brazilian inequality is not falling uniformly. While the whole country has shown a constant reduction on inequality in the first decade of the twenty-first century, on the Metropolitan Areas we observe a persistent income disparity. The interior of the States, on the opposite, have presented a consistent fall on inequality of the household per capita income and wages. These results suggest that a great part of inequality reduction in Brazil was related to the shrinking differences between regions and between Metropolises and interiors, following Azzoni and Neto (2011) and Souza and Osório (2011).

In order to understand why inequality has not been falling on big cities, it is necessary more studies investigating their particularities. All of these cities have presented a reduction on the inequality of wages, but it was less intense than in the rest of the country. This reduction is a consequence of the higher education of the population, the constant increase on the value of the minimum wage together with a trend of formalization of the workforce. With a minor impact, but still important, there is less discrimination of gender and race than before (Barros et al, 2007; Ferreira et al, 2006). Among the ten biggest Brazilian municipalities, Brasília had the smallest reduction on wages inequality. The results exposed here justify this pattern, as long as the elevation on public servants' wages perpetuated the disparities on the federal capital.

Besides the public sector, the slightly rise on the inequality of the household per capita income is also caused by the higher inequality of income from other sources and the qualified immigration to Brasília. However, these two factors can also be associated to the public sector. The income from other sources has probably been more unequal in part because of pensions, whose value is much higher in the public sector (follows the salary). Also, as demonstrated, a great part of the qualified immigrants consists of public servants attracted by the high wages in Brasília. The Censuses data do not allow for the quantification of those two factors on inequality, nevertheless, if it was possible, the effect of the public sector on total inequality would be even greater.

Another topic that could affect the household per capita income inequality is the family size. As put above, this variable has not had relevance to explain the behavior of inequality on the Federal District. However, we have not tested the profile of the families on the two years (people who works, marriages, children…). Other studies can investigate whether marriages between members of the same social class and the insertion of siblings on

the labor force are an important factor to explain the maintenance of high levels of inequality on the Brazilian metropolises.

All the methods employed have led to a common result. The inequality on the Federal District has risen on those 10 years. The main force causing this phenomenon is the State, basically through the remuneration of its employees. The government must be aware of the causes of its acts on inequality. In Brasilia, this effect is evident, but it can also have some relevance in other big cities.

Facing it, the governments have a difficult role: to conciliate the attractiveness of the jobs on the public sector with the possible negative effects of it. The governments ought to resist the pressure for better wages, and, at the same time, keep the public sector attractive for new qualified employees. Policies of better wages make the public sector more attractive, but also make the budget more restrict and, as shown here, can increase inequality at least locally. In addition, it is also not clear if higher wages on the public sector make it more efficient (Holanda, 2009; Heitmueller, 2006). Those wages can allocate on the public sector people without the profile for it. This would be a sign that this sector is attracting more people than it would on the efficient equilibrium. This pattern would create distortions. In Brasilia, the idea of being a public servant as the ideal job is continuously more observed, in detriment of other positions or entrepreneurship. Besides the long run impact on the public budget (by the rigid legislation and the pensions), we have proved here that the income distribution is more concentrated under these conditions.

The use of the Shapley value with Brazilian data is consistent and rigorous. New studies about the Brazilian inequality must be done and this method has a great potential. These studies can find answers to the different trends of income distribution on the country. The great cities seem to suffer from a clear dichotomy between rich people and poor people, which continues since the migration from rural areas to urban ones during the middle of last century. The recent fall of inequality has not happened on the metropolises. Subsidize new policies to revert this pattern is a challenge for researches.

## References


ARAAR, A; DUCLOS, J-Y. Distributive Analysis Stata Package - DASP. User Manual, DASP v.2.2. 2007.

AZZONI, C. R.; NETO, R. M. S. Non-spatial government policies and regional income inequality in Brazil. *Regional Studies*, v. 45. 2011.

BARROS, R. P.; FRANCO, S; MENDONÇA, R. A recente queda da desigualdade de renda e o acelerado progresso educacional brasileiro da última década. Texto para discussão nº 1304, Instituto de Pesquisa Econômica Aplicada - IPEA. Brasília. 2007.

BAYE, F.; EPO, B. Inequality Decomposition by regressed-income sources in Cameroon. In: *Annals of the Conference on Measuring National Income, Wealth, Poverty and Inequality in African Countries*. Cape Town, September 2011.

BECKER. G. S.; PHILIPSON, T. J. e SOARES, R. R. The quantity and quality of life and the evolution of World inequality. *American Economic Review*, v. 95. 2005.

BELLÙ, L. G.; LIBERATI, P; Decomposition of Income Inequality by Subgroups. EASYPol – On-line resource materials for policy making. Food and Agriculture Organization of the United Nations. 2006. Available in: www.fao.org/docs/up/easypol/444/dcmpsng-inqulty_sbgrp_052en.pdf

BENDER, S.; FERNANDES, R. Gastos Públicos com pessoal: Uma análise de emprego e salário no setor público brasileiro nos anos 90. *Revista EconomiA*, v.10. 2009.

BOURGUIGNON, F. Decomposable Income Distribution Measures. *Econometrica*, v. 47. 1979.

BOURGUIGNON, F.; MORRISON, C. Inequality among World Citizens: 1820-1992. *American Economic Review*, v. 92. 2002.

CHARPENTIER, A; MUSSARD, S. Income Inequality Games. Cahier n° 2010-05, Ecole Polytechnique, Centre National de la Recherche Scientifique. 2010.

COWELL, F.; JENKINS, S. How much inequality can we explain? A methodology and an application to the United States. *The Economic Journal*, v. 105. 1995.

COWELL, F. Measuring Inequality. Oxford University Press, 1st edition, London, 2009.

DA MATA, D. ET AL. Quais características das cidades determinam a atração de migrantes qualificados? Texto para discussão nº 1305, IPEA. 2007.

DATASUS. Ministério da Saúde (Ministry of Health). *Estatísticas Demográficas e Socioeconômicas*. Available in: www.datasus.gov.br/DATASUS/index.php?area=0206.

EBLE, A. Decomposing Household Income by Source and Subgroup: A Methodological Investigation Using Survey Data from Rural China. Indiana University, mimeo. 2007.

FIELDS, G. Accounting for Income Inequality and its Change: A New Method, With Application to the Distribution of Earnings in the United States. First Version: Cornell University, mimeo. 1998. Final version: *Research in Labor Economics*, v. 22. 2003.

FIELDS, G.; YOO, G. Falling Labor Income Inequality in Korea's Economic Growth: Patterns and Underlying Causes, *Review of Income and Wealth*, v. 46. 2000.

FERREIRA F.; LEITE, P.; LITCHFIELD, J.; ULYSSEA, G. Ascensão e queda da desigualdade de renda no Brasil. *Econômica*, v. 8. 2006.

GARCIA, A. S. Desigualdades raciais e segregação urbana em antigas capitais. PhD Thesis. Universidade Federal do Rio de Janeiro – Federal University of Rio de Janeiro. Rio de Janeiro. 2006.

GOMES, D. C.; SILVA, L; SÓRIA, S. Condições e relações de trabalho no serviço público: o caso do governo Lula. *Revista de Sociologia e Política*, v. 20. 2012.

GUSTAFSOON, B.; JOHANSSON, M. In search of smoking guns: what makes income inequality vary over time in different countries? *American Sociological Review*, v. 64. 1999.

HEITMUELLER, A. Public-private sector pay differentials in a devolved Scotland. *Journal of Applied Economics*, v. 9. 2006.

HOLANDA, A. L.; Diferencial de Salários entre os setores público e privado: uma resenha da literatura. Texto para discussão nº 1457, IPEA. 2009.

HOLANDA, A. L.; BARBOSA, F. Diferencial de Salários entre os setores público e privado no Brasil: um modelo de escolha endógena. Texto para discussão nº 1753, IPEA. 2010.

IBGE. Instituto Brasileiro de Geografia e Estatística. Censo 2000. Brasília, 2002. Available in: www.ibge.gov.br/censo/divulgacao.sht

_____. Censo 2010. Brasília, 2012. Available in: www.ibge.gov.br/censo2010

_____. Características da População e dos Domicílios. Censo Demográfico 2010, resultados do universo. Rio de Janeiro. 2011.

LITCHFIELD, J. Welfare and Income Distribution in Brazil, 1981-1995. Ph.D. Thesis. London School of Economics. 2001.

MINCER, J. *Schooling, Experience and Earnings*. Columbia University Press, New York, 1974.

MINISTÉRIO DO PLANEJAMENTO E GESTÃO (MPOG). *Boletim Estatístico de Pessoal*, v. 15, nº 176. Brasília, Dez. 2010.

MORDUCH, J; SICULAR, T. Rethinking inequality decomposition with evidence from rural

China. *The Economic Journal*, v.112. 2002.

OAXACA, R. Male-female wage differentials in urban labor markets. *International Economic Review*, v. 14, 1973.

SECRETARIA DA TRANSPARÊNCIA E CONTROLE DO DF. Portal da Transparência. 2013. Available in: www.transparencia.df.gov.br/SitePages/Portal%20Da%20Transpar%C3%AAncia.aspx

SALARDI, P. How much of Brazilian inequality can we explain? Catholic University of Piacenza, mimeo. 2005.

SANTOS, C.; FERREIRA, P. C.; Migração e distribuição regional de renda no Brasil. *Pesquisa e Planejamento Econômico*, v.37. 2007.

SASTRE, M; TRANNOY, A. Shapley inequality decomposition by factor components: some methodological issues. *Journal of Economics*, v. 9. 2002.

SHAPLEY, L. S. A Value for *n*-person Games. In: *Contributions to the Theory of Games*, volume II, by H.W. Kuhn and A.W. Tucker, editors. *Annals of Mathematical Studies* v. 28, pp. 307–317. Princeton University Press, 1953.

SHORROCKS, A. The Class of Additively Decomposable Inequality Measures. *Econometrica* v. 48. 1982.

_____, A. Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. First version: University of Essex, mimeo. 1999. Final version: *Journal of Economic Inequality*, v. 11. 2012.

SOARES. R. R. Health and the evolution of welfare across Brazilian municipalities. *Journal of Development Economics*, v. 84. 2007.

SOARES, S.; OSÓRIO, R.; VERAS, F.; MEDEIROS, M.; ZEPEDA, E. Programas de Transferência Condicionada de Renda no Brasil, Chile e México: Impactos Sobre a Desigualdade, Texto para discussão nº 1293, IPEA. 2007.

SOUZA, P. A distribuição de renda nas pesquisas domiciliares brasileiras: harmonização e comparação entre Censos, PNADs e POFs. Texto para discussão nº 1832. IPEA. 2013.

SOUZA, P.; OSÓRIO, R. A redução das disparidades regionais e a queda da desigualdade nacional de renda (1981-2009). Texto para discussão nº 1648, IPEA. 2011.

SOUZA, P.; MEDEIROS, M. Diferencial salarial público-privado e desigualdade de renda *per capita* no Brasil. *Estudos Econômicos*, vol. 43. 2013a.

_____. Gastos Públicos, tributos e desigualdade de renda no Brasil. Texto para discussão nº 1844, IPEA. 2013b.

UNITED NATIONS HABITAT. State of the World's Cities 2010/2011 – Bridging the Urban Divide. London. 2008.

VAZ, D; HOFFMANN, R. Remuneração nos serviços no Brasil: o contraste entre funcionários públicos e privados. *Economia e Sociedade*, v. 16. 2007.

WAN, G. Regression-based inequality decomposition: pitfalls and a solution procedure. Discussion Paper nº 2002/101. United Nations University – WIDER. Helsinki, 2002.

_____. Accounting for income inequality in rural China: a regression based approach. *Journal of Comparative Economics*, v. 32. 2004.

WAN, G; ZHOU, Z. Income inequality in rural China: Regression based decomposition using household data. Research Paper nº 2004/51. United Nations University – WIDER. Helsinki, 2004.

WOOLDRIDGE, Jeffrey. Econometric analysis of cross section and panel data. 1st Edition. Cambridge, Massachusetts. MIT Press. 2002.

## Appendix A – Database used: Brazilian Censuses of 2000 and 2010

We have used the Census of 2000 and 2010 made by the 'Instituto Brasileiro de Geografia e Estatística (IBGE)' – Brazilian Institute of Geography and Statistics. We used the data of people and households. In order to measure the inequality in the biggest Brazilian cities, we used the whole Census. But, on the other sections of the text, the data used was restricted to Brasilia, the Federal District of Brazil.

The Census data are collected every ten years by the Institute. It has two aims: count the population of the country and interview a part of this population in order to obtain a disaggregated socio-demographic portrait of the country and its municipalities. In this interview, questions about income are included, differently from the majority of the Census in other countries. Since the complex structure of the sample, each household has a parameter of expansion (how many households the interviewed one represents), or, in other words, a different weight.

In 2010, the sample size for the Federal District was 200,888 in 2000 and 116,458 in 2010. Because there were changes on sampling methods, we have seen a reduction on the size of the sample in 2010. Thus, each household has been more representative in the latter year. We believe the samples of both years are robust, even after its shrinking, given the reputation of IBGE at conducting the national surveys.

For a direct comparison between 2000 and 2010, we decided to keep the spatial configuration of 2000. The Federal District had nineteen Administrative Regions in 2000, which were added by eleven new ones during this time. The new regions were simply aggregated into their previous localities. Also, the Census distinguish between urban and rural areas. Thus, we opted to treat the rural area as a region, because of its singularities and small population. In the end, we had twenty regions, 19 with the urban areas of the Administrative Regions of 2000 and one corresponding to the rural area of the Federal District.

# Appendix B – Wages differentials between public servants and private sector workers on the formal labor market of the Federal District of Brazil

## 1. Estimating the determinants of the hourly wage

The regression of the determinants of wages on the formal sector allows us to compare directly the disparities on the remuneration of public servants and its equivalent group, the private sector workers in the formal market. The analysis resembles the ones made by Vaz and Hoffman (2007), Bender and Fernandes (2009) and Souza and Medeiros (2013a) for Brazil. In order to have a more homogeneous group of workers, we selected the individuals with 25 years old or more. We suppose that with this age, people have already finished their studies and are on the labor market. On this analysis, we take into account only the remuneration of the main job of persons who are public servants or formal private sector workers. We do not include in the regression the domestic or informal workers, as well as the employers or unoccupied.

The public servants corresponded to 10.6% of the total occupied population above 25 years old in 2000, rising to 13.4% in 2010. The formal private market had 31.3% in 2000 and 49.9% in 2010. With the sample restricted to the formal workers with 25 years old or more, the public servants were 28.4% in 2000 and 23.2% in 2010.

Examining the Census of 2000 and 2010, we verified that the mean monthly wage of the public servants has climbed from R$ 4,278 in 2000 to R$ 6,246 in 2010 (real increase of 46%). Considering the hourly wage, it was R$ 28.24 in 2000 and R$ 47.60 in 2010, 68.5% more. The mean monthly wage of the formal private sector workers had contracted slightly, from R$ 2,354 to R$ 2,198. Informal workers with lower incomes started to go to the formal sector during this decade, which has reduced the average income of this group. Nevertheless, the hourly wage of this group has had an increase from R$ 15.15 to R$ 18.40[13].

Clearly, the wage difference between public and private workers has risen during these 10 years. The mean wage of the public servants was equivalent to 181.7% of the mean of formal private sector workers in 2000, climbing to 284.1% in 2010 (from 186.4% to

[13] Other studies can explore why the monthly wage has dropped while the hourly wage has elevated in the Federal District. The data suggest that people would be supplying less work even when faced with higher incomes. It can be analyzed on future researches – possible rise on the number of part-time jobs, introduction of more women into the labor force, among others.

258.7% when it comes to the hourly wage). This rise on the gross difference of the wages is not conclusive, as it can come from divergences in the attributes of the two groups in the two moments of time. On Table 8, we can see the educational distribution of the two groups, as well as the whole population with 25 years old or more. This table shows that, on average, the public servants are more educated than the others, which explains part of their better wages.

**Table 8 – Education of the population with 25 years or more by Type of Occupation**

| | Formal Private Sector Worker | | Public Servant | | People with 25 years or more | |
|---|---|---|---|---|---|---|
| | **2000** | **2010** | **2000** | **2010** | **2000** | **2010** |
| **Illiterate** | 2.76% | 2.63% | 0.47% | 0.44% | 6.98% | 4.52% |
| **Primary School** | 37.28% | 37.57% | 12.21% | 9.54% | 44.45% | 40.08% |
| **High School** | 34.68% | 36.33% | 35.78% | 27.91% | 27.99% | 31.45% |
| **College or more** | 25.28% | 23.47% | 51.54% | 62.10% | 20.58% | 23.96% |

Source: Censuses 2000 and 2010, IBGE. Author's calculation.

The drop on the proportion of workers with college degree in the formal private sector and the rise on the number of workers with primary education is certainly a consequence of the great formalization of the Brazilian labor market during the decade. People who used to work in domestic or informal jobs started to get into the formal sector of the economy. It is not a surprised that this people incorporated to the formal sector present a lower level of education. On the opposite side, the public servants had more college degree holders in 2010 than in 2000. The percentage has changed from 51.5% to 62.1%, which demonstrates a greater demand from the government for qualified professionals.

As the simply comparison of wages is not conclusive, we used a regression approach. We estimate an equation of the determinants of the wages on the formal sector of the Federal District (public and private ones). The regressions have the same structure for 2000 and 2010, following Bender and Fernandes (2009). The independent variable is the logarithmic of the hourly wage. We opted to use the hourly wage as it yields the real remuneration of the work effectively supplied. Differently from the model on the main text, we used experience instead of age. We also added a quadratic term for this variable. It fits better the theory about labor market returns, as the wages have a decreasing trend after a certain time on the labor force – human capital depreciation (Mincer, 1974). The individual experience is defined by his age subtracted of the six first years of his life and the years of schooling completed. Thus, *Experience = Age – (Education + 6)*. The estimated equation is:

$$Ln\ (Y_{sh}) = \beta_0 + \beta_1\ Education + \beta_2\ Experience + \beta_3\ Experience\ Squared + \beta_4\ White + \beta_5\ Woman + \beta_6\ Public\ Servant + \beta_7\ Rich\ Region + \beta_8\ Middle\ Income\ Region + \varepsilon$$

The variables are similar to the regression on the main text. *Public Servant* is a dummy that yields the wage difference in relation to the base group, the formal private sector workers. The results are on Table 9 below:

**Table 9 – Determinants of the hourly wage on the formal labor market Brasilia - Federal District, 2000 and 2010. Individuals with 25 years old or more. Dependent Variable: Logarithmic of hourly wage**

| | 2000 | 2010 |
|---|---|---|
| **Intercept** | -0.65934 | 0.36247 |
| **Education** | 0.14665 | 0.12734 |
| **Experience** | 0.03758 | 0.025716 |
| **Experience Squared** | -0.00027 | -0.00008 |
| **White** | 0.1007 | 0.13673 |
| **Woman** | -0.25754 | -0.25554 |
| **Public Servant** | 0.39526 | 0.61286 |
| **Rich Region** | 0.55156 | 0.63375 |
| **Middle Income Region** | 0.23184 | 0.27293 |
| **F Test** | 6,944 | 3,542 |
| **$R^2$** | 0.603 | 0.528 |
| **Observations** | 36,624 | 25,305 |

Note 1: All the estimations were significant at the 1% level.
Note 2: We have used the method of Weighted Least Squares, on which each observation has a weight attributed by IBGE.

According to the estimations, there were a reduction on the returns of education and experience. Each additional year of schooling yielded a return on wages of 14.6% in 2000 and 12.7% in 2010. The contribution of experience has also decreased. For the case of an employee with 10 years of experience, his salary was 35% greater in 2000 compared to

another employee without any experience. In 2010, this difference shrank to 25%. The drop happens for any level of experience. The wage disparity between whites and non-whites has risen, from 10% to 13.6%. This change is probably a consequence of the greater proportion of people declaring themselves blacks or mixed[14]. We have not seen a considerable change on the difference between men and women, as it stayed around 25% on both years. The disparities of wages between regions have increased during the decade. In opposite, as we have seen on the regression for the household per capita income, this difference has shortened. These differing results are probably connected to the great formalization during the years. The formalization was most likely to occur on the poor regions, which made the wages of the formal sector more unequal between regions, even though the regional disparities of the total income have diminished.

The variable whose impact changes the most between 2000 and 2010 was *Public servant*. In 2000, this group had an average wage 39.5% better than the private sector one, *ceteris paribus*. In 2010, this difference has climbed to 61.3%. As we are controlling for other variables, we can notice that there were a real increase on the wages of the former group, overwhelming the latter group. The changing attributes of the groups was not the only factor which has raised the wage differential. There was a continuous improvement of the public sector conditions of working, which made this sector more attractive. This result follows Bender and Fernandes (2009) and Souza and Medeiros (2013a) for Brazil. As Brasilia is the center of the federal public administration, and the government of the district pays the best salaries among all the Brazilian federation units, we could expect a higher wage premium for the public servants of Brasilia than in the other cities.

On the next section, we make the regression-based decomposition of wages inequality in the two years under analyses. We can see that this better wage premium has a strong effect on the inequality of the formal labor market of the Federal District of Brazil.

[14] Perhaps, the people who declared themselves whites in 2000 and started to declare themselves blacks or mixed in 2010 have an average income smaller than the ones who continued to declare themselves whites. This pattern would push the increase on the average income of whites in comparison to non-whites. This is a topic which can be investigated in other studies.

## 2. Decomposition of Inequality

On Table 10 below, there is the decomposition of inequality of the hourly wage on the formal labor market of Brasilia. It yields us the contribution of each variable to total inequality in 2000 and 2010. The previous regression is the base for it. The results are complements the ones in the main text and show the same pattern.

**Table 10 – Regression-based Decomposition of Inequality - Formal sector labor market. Brasilia – Federal District of Brazil, 2000 and 2010. Individuals with 25 years or more.**

| | Fields | Shapley Value | | | | | |
|---|---|---|---|---|---|---|---|
| | % | Gini | % | GE(0) | % | GE(1) | % |
| **2000** | | | | | | | |
| Education | 35.73 | 0.1736 | 32.33 | 0.2042 | 38.54 | 0.1709 | 32.68 |
| Experience | 4.93 | 0.0559 | 10.41 | 0.0055 | 1.04 | 0.0205 | 3.92 |
| Race/color | 1.34 | 0.0084 | 1.56 | 0.0071 | 1.34 | 0.0066 | 1.26 |
| Gender | 0.10 | 0.0126 | 2.35 | 0.0027 | 0.51 | 0.0055 | 1.05 |
| Public Servants | 6.60 | 0.0306 | 5.70 | 0.0289 | 5.46 | 0.0218 | 4.17 |
| Region | 11.58 | 0.0627 | 11.68 | 0.0605 | 11.42 | 0.0550 | 10.52 |
| Residual | 39.71 | 0.1929 | 35.96 | 0.2207 | 41.66 | 0.2426 | 46.39 |
| **Total** | **100** | **0.537** | **100** | **0.530** | **100** | **0.523** | **100** |
| **2010** | | | | | | | |
| Education | 24.01 | 0.132 | 21.26 | 0.1605 | 21.89 | 0.0970 | 11.13 |
| Experience | 5.01 | 0.0507 | 8.16 | 0.0255 | 3.48 | 0.0724 | 8.31 |
| Race/color | 1.61 | 0.0118 | 1.90 | 0.011 | 1.50 | 0.0113 | 1.30 |
| Gender | 0.77 | 0.0129 | 2.08 | 0.0026 | 0.35 | -0.0045 | -0.52 |
| Public Servants | 11.19 | 0.0582 | 9.37 | 0.0643 | 8.77 | 0.0444 | 5.09 |
| Region | 10.63 | 0.0701 | 11.29 | 0.0714 | 9.74 | 0.0692 | 7.94 |
| Residual | 46.80 | 0.2856 | 45.99 | 0.3976 | 54.23 | 0.5818 | 66.75 |
| **Total** | **100** | **0.621** | **100** | **0.7332** | **100** | **0.8716** | **100** |

Source: Censuses 2000 and 2010, IBGE. Author's calculation.

Similarly to the analysis made for the household per capita income, the contribution of education on the hourly wage inequality has also reduced between 2000 and 2010. But, the drop shown here was even stronger. The proportion explained by this variable was at least 32% of total inequality in 2000, falling to less than 22% in 2010. In spite of the elevation of inequality in the formal labor market, the contribution of education has fallen proportionally

and in total, does not matter the index or method used. The experience has not shown a strong pattern, as it has a smaller contribution measured by the Gini Index but a bigger one by other methods. Altogether, there is a rise of 10% on the residual term. This probably is a consequence of the formalization, which has made the inequality harder to be explained and described with the same variables.

Differences of gender and race/color have also fallen slightly, even with the regressions indicating greater returns for whites in 2010. Nevertheless, those variables explain a little portion of total inequality, demonstrating that homogenizing those groups would have a small impact on the wages distribution.

The regional differences also had its impact on inequality diminished. But, as before, we cannot conclude if there were a reduction on regional disparities or it was caused by the aggregation we have made.

Finally, as in the case of the household per capita income, the only variable which has had a significant increase on its contribution to total inequality was the dummy for public servants. It has happened in any method used. By the Shapley value using the Gini index, the contribution of this variable has climbed, from 5.70% to 9.37%. The inequality on the formal sector labor market has risen as well as the relative contribution of the public servants. It shows that the absolute contribution of this variable have risen even more intensively, from 0.0306 to 0.0582, using the Gini index. If the absolute contribution of this variable in 2010 was the same of 2000, the measured Gini index would be 0.593 instead of 0.621, 4.5% less. We can conclude that the elevation of inequality in the formal labor market was not only caused by the great formalization in the period, but a considerable part of it was a consequence of the wages paid by the public sector.

## Appendix B References


BENDER, S.; FERNANDES, R. Gastos Públicos com pessoal: Uma análise de emprego e salário no setor público brasileiro nos anos 90. *Revista EconomiA*, v.10. 2009.

SOUZA, P.; MEDEIROS, M. Diferencial salarial público-privado e desigualdade de renda *per capita* no Brasil. *Estudos Econômicos*, vol. 43. 2013a.

VAZ, D; HOFFMANN, R. Remuneração nos serviços no Brasil: o contraste entre funcionários públicos e privados. *Economia e Sociedade*, v. 16. 2007.