New Insights About Wage Inequality in Colombia

Andres Gomez-Lievano, Juan Tellez
and Eduardo Lora
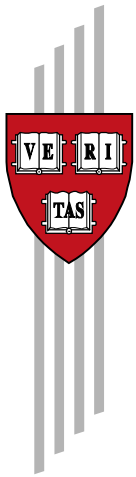
Working Papers

Center for International Development
at Harvard University

New Insights About Wage Inequality in Colombia

Andres Gomez-Lievano[1],[*], Juan Tellez[1],[*], and Eduardo Lora[1]

(1) Harvard Kennedy School, Harvard University, Cambridge, MA

(*) Both authors contributed equally to this work.

Abstract

This paper presents a descriptive analysis of wage inequality in Colombia by cities and industries and attempts to evaluate the impact of the inequality of industries on inequality of cities. Using the 2104 Colombian Social Security data, we calculate the gini coefficient for cities and industries and draw comparisons between their distributions. Our results show that while cities are unequal in similar ways, industries differ widely on how unequal they can be with ginis. Moreover, industrial structure plays a significant role to determine city inequality. Industrial framework proves to be a key element in this area for researches and policymakers.

1   Introduction

Inequality has undoubtedly become one of the most crucial challenges of the contemporary era. In general, the main purpose of studying the inequality phenomenon is to provide a picture of the distribution of income (or welfare) disparities among several individuals, in order to take the required policy measures to correct them if it is necessary. In a 2015 publication, The Organization of Economic Co-operation and Development (OECD) showed that the poorest 10 % people earn 9.6 times less the income of the richest 10 % in the OECD countries (OECD, 2015). The Credit Suisse Research Institute and The Oxford Committee for Famine Relief (OXFAM) estimated in 2014 that the 48% of global wealth (net worth in financial and real assets) is held by the richest 1% (Credit Suisse Research Institute, 2014). All things being equal, it is expected than in 2020 this proportion will increase to 54%.

In Latin America, income inequality has been decreasing since the 1980s but the levels remain high compared to other regions of the world (SEDLAC). Colombia experienced a period of sustainable growth from 2002 until

2012, registering in 2013 the lowest gini value, 53.5, since 1995 (World Development Indicators). But even with this improvement, Colombia still occupies the 127th place out of 138 countries in terms of income inequality, which shows that there is still a need for policies that address this issue. With the goal of generating significant advances for a more equal society, the government created the programs 2014-2018 National Development Plan (NDP) "Todos por un nuevo país", and 2010-2014 NDP "Prosperidad para todos", that achieved positive results.

Nevertheless, government intervention has not always been welcomed. It has longtime been argued that the role of income equality had a mixed effect on economic growth. This view has changed in the last decade, where institutions such as the International Monetary Fund (IMF) and the World Economic Forum (WEF) have stressed on the key role of equality in order to achieve sustainable growth. For example, Berg and Ostry (2011) found evidence that a more equal society fosters long term growth, being a factor at least as important as free trade, political institutions and macroeconomic stability among others. Consequently, the reduction of inequality has become an important target for policymakers to the extent of start adopting the term *inclusive growth* more than only growth.

Essentially, income is divided into two sources: capital gains and labor income. Each contributes differently to inequality (Piketty, 2014). Gains from capital are difficult to study because tax agencies do not typically share these data for research purposes. Hence, gains from capital are restricted to household surveys. In contrast, social security datasets have become increasingly available to researchers, providing the opportunity to analyze the full population of formal workers. Here, we study the wage inequality using the 2014 Colombian Social Security data that contains all formal workers in the country. While this data has a number of limitations, particularly because in Colombia only 30% of the working age population contributes to the social security. However, access to these data is an opportunity to investigate in full depth the wages of all formal workers across several years. In addition, the study by Pineda and Acosta (2009) shows that at a national level labor income represents approximately 80% of income in Colombia.

The literature at large has studied wage inequality in the context of nations, inequality due differences in education and years of experience. But few studies have used the fact that labor can be differentiated in a variety of ways, and is expressed in a variety of economic activities. Our contribution is to analyze wage inequality in Colombia through the formal sector, and describing the role that industries have in general, and in the different

regions.

The main objective of this paper is twofold. We want to show an accurate portrait of the most recent state of the wage inequality in Colombia. Second, we would like to assess the impact of industrial structure on wage inequality of cities, by using a linear regression model and the same data. This would allow us to know to what extent inequality of industries present in a particular city can provide any information about the inequality of this city.

The papers is organized as follows. The next section presents a literature review on inequality. Section 3 describes our database, explaining how the wage variable was calculated. Section 4 is divided in two subsections, the first shows the results of our analysis of wage inequality by cities and the second subsection by industries. Section 5 presents the model and the estimation results. The final section draws conclusions and suggests further research steps.

2   Literature Review

Colombia has mainly been focused in understanding inequality in different regions (Bonilla, 2008; Garza, 2008). We go beyond regions and also analyze inequality by economic activity. In the United States this type of analysis has been more popular, thanks to a more institutionalized interest in inequality and public policy (Galbraith and Hale, 2014; Moore, 2009). In Latin America, Marshall (2010) and Esquivel (2011) study income inequality by industries in Argentina and Mexico respectively. While those mentioned papers consider all sectors in their analysis, they both focus on manufacturing.

Aside from a geographical study of the wage inequality in Colombia, others perspectives have mostly been centered in explaining the factors that contribute to it; such as education, experience and gender. One of the goals of Bonilla (2009) is to identify the determinants and to what extent they affect income inequality by regions; putting an emphasis on the level of education. Pineda and Acosta (2009) analyzes income distribution and labor market in Colombia, focusing on a a gender perspective. In their case, they study wages as well as capital gains.

This paper is the first attempt to draw a common picture, compare and establish a relationship of the income inequality distribution of cities and industries in Colombia. This could prove to be helpful and insightful in future policymaking processes.

3 Data

We use the Social Security database of Colombia as the source of our analysis on wages. This database consists of all the monthly contributions to the social security system that firms pay to their workers. Hence, our population of analysis is the set of formal workers who contribute to the social security in 2014.

Our database also reports the municipality the worker lives and the industry he or she works in. Industries are classified according to the ISIC Revision 3.0 classification, for which there are a total of 445 unique 4-digit industry codes. We use Duranton (2013) methodology to delineate 19 metropolitan areas that consist of a collection of municipalities that are strongly connected in terms of commuting ties. This definition is an attempt to delineate labor markets. We consider 43 additional municipalities as labor markets that have populations above 50,000 and are 75% urban. Hence, our geographical units of analysis are 62 cities.

We aggregate this data in a way such that, for each worker, we calculate the net wage earned in each year, we divide by the total days of work that were reported in the social security system, and multiply by 30 to get the average monthly wage of the worker. Table 1 shows some descriptive statistics.

Table 1: Descriptives across year of our social security dataset in Colombia. Wages are given in Colombian Pesos.

Year	Mean Wage	Median Wage	Top 1% Wage	Number of workers
2008	$797,562	$461,633	$5,323,333	9,847,641
2009	$845,885	$497,000	$5,472,197	10,443,018
2010	$901,460	$515,000	$6,171,840	11,288,276
2011	$931,411	$535,645	$6,138,909	12,435,433
2012	$993,191	$567,000	$6,500,000	13,157,556
2013	$1,038,234	$589,500	$6,832,224	13,651,794
2014	$1,061,225	$616,000	$6,702,310	15,255,258

4 Descriptive Analysis

Figure 1 shows the boxplots of the gini coefficients for all cities and industries. Both distributions differ in several aspects. First, while both cities and industries have ginis that are approximately symmetrical, they are around 28% for cities and around 33% for industries, and this difference is statistically significant ($t = -4.37$). Second, the standard deviation of ginis for industries is twice as for cities. And third, cities have a tendency to have some

outliers, whereas industries have ginis more or less uniformly distributed (approximately between 0.5% and 65%. This indicates that there is an important proportion of significantly *smaller* gini values for some cities.
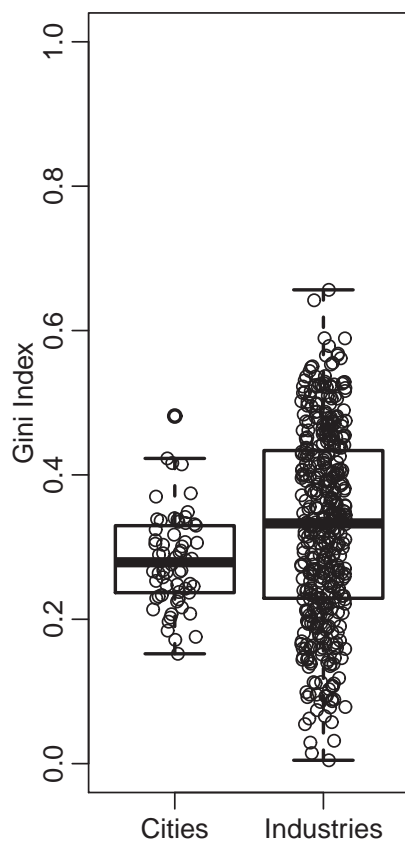

Figure 1: Boxplot of ginis across cities and industries in Colombia in 2014 (only formal employment).

Given that industries have a much higher gini variability than cities, this leads us to two different scenarios: industries can be much more unequal than cities but also much more equal. It is interesting to note that the first quartile is almost the same for both distributions, and then the second and third quartiles are higher for industries. The behavior is different, however, since there are no outliers for industries but there are some for cities.

By taking the industries boxplot as a benchmark, the most relevant fact that we can retain from this figure is that cities and industries seem to have different dynamics. It is safe to assume cities do not appear to be

disproportionately concentrating industries that have similar inequalities (either on the high or low end of ginis). Instead, cities seem to be mixing the inequality of industries.

4.1 Inequality across cities

Most economic activity happens in cities, almost by definition. This fact, however, makes cities very unequal places. Rich get richer, differences generate segregation, and the pattern repeats itself. The profile of inequality in Colombian cities is not surprising. Bigger cities are more unequal, and gini indices cluster around 0.3 (see Figure 1).

Table 2 and Table 3 show gini coefficients for the five least and most unequal cities in Colombia in 2014. The city in Colombia that has the most equal income distribution (in the formal sector) is Apartadó, and the most unequal distribution is Montelíbano.

Table 2: Least unequal cities in Colombia in 2014

City	Gini in %
Apartadó	15.2
Cartago	17.1
Chigorodó	17.5
Girardot Met	18.4
El Carmen del Bolívar	19.6

Table 3: Most unequal cities in Colombia in 2014

City	Gini in %
Cartagena Met	37.4
Barrancamermeja	41.4
Arauca	41.6
Bogotá	42.2
Montelíbano	48.1

As we mentioned, we want to go beyond cities as the unit of analysis, and also study industries. Since people do not just work in cities, but also in specific industries, understanding the relationship between the inequality

in cities and industries can provide insights about how to propose economic growth that can be more inclusive.

4.2 Inequality across industries

Table 4 shows the five least unequal industries in Colombia for the year 2014. The first four, which are the most equal industries, belong to the *manufacturing* sector. This result is suggesting a link between income equality and economic activity. The remaining industry in this top five belong to the *wholesale and retail trade* sector. Let us notice that, besides being at the top of the equality distribution; these aforementioned industries have very low gini coefficients (less than 6 %).

Table 4: Least unequal industries in Colombia in 2014

Code	Industry	Gini in %
1924	Manufacture of plastic footwear with any type of sole, except sports footwear	0.4
2927	Manufacture of weapons and ammunition	1.4
3692	Manufacture of musical instruments	2.9
1925	Manufacture of sports footwear, including molded	3.1
5112	Wholesale on a fee or contract basis of parchment coffee	5.4

Table 5 shows the five most unequal industries in Colombia for the year 2014. This time the situation is different from the least unequal industries, if we group them by sectors. Two of them belong to *transport, storage and communications*, two others to *mining and quarrying* and one to *financial intermediation*. It is interesting to notice that no industry of these top five unequal is in the manufacturing category, which gives more evidence for a possible relationship of income inequality and some economic activities.

Table 5: Most unequal industries in Colombia in 2014

Code	Industry	Gini in %
6120	Inland water transport	57.8
1331	Mining of nickel ores	58.9
1030	Extraction and agglomeration of peat	58.9
6714	Other activities related to the stock market	64.1
6050	Transport via pipelines	65.6

5 Do Economic Activities Affect Urban Inequality?

As we have already mentioned, the observation that the variability of industry ginis is much wider than the variability of city ginis suggests that cities have an averaging effect reducing the inequalities that characterize each industry. We test this by constructing, for each city, an average gini from the ginis of industries that are present, weighted by the number of workers that are employed in each industry. Hence, we have $X_c = \sum_i gini_i * (E_{c,i}/E_c)$, where $E_{c,i}$ is the number of workers in city c employed in industry i, and E_c is the total size of formal employment in city c.
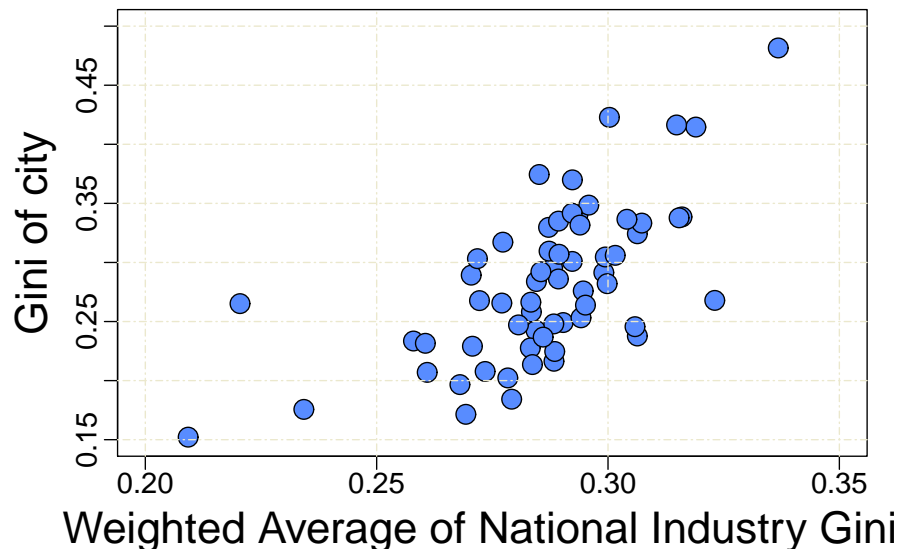

Figure 2: Plot of the gini of cities against weighted average of national industry ginis.

Figure 2 plots the real gini of cities against our prediction from an averaging of industry ginis, in the year 2014. The correlation is $r = 0.64$, meaning that $R^2 = 0.41$. It is important to emphasize that the industry ginis are national. This means that 40% of the variability in ginis across cities can be explained by the *within* industry inequality of the industries in a place. Even if we find a high correlation, our result reveals that there is a portion of the wage inequality in cities associated to inequalities *between* industries.

The converse analysis can be done for industries. We calculate in this way an average of city ginis for each industry, according to how many employees it has across cities, such that $Y_i = \sum_c gini_c * (E_{c,i}/E_i)$, where E_i is the total size of formal employment in industry i. This weighted average gini Y_i only explains 4% of the variation of

real gini industries.

6   Conclusion

This paper aimed to provide two contributions to the inequality literature in Colombia, with the motivation of bringing out insightful elements to generate accurate economic and social policies. First, we characterized the present state of inequality in the formal economy in Colombia, using gini indices for cities and industries. For this we used the 2014 Social Security data. Second, we were able to elucidate information about how industry inequality affect city inequality.

Industries are very heterogeneous in terms of income inequality, meaning that industries, despite sharing within Colombia similar institutions, laws, and prices, differ widely in the way the distribute their profits among their workers. Cities, in contrast, are much more homogenous, suggesting that industries with different income inequalities tend to be present together. This means that geographically speaking there is a tendency of places to *average out industry inequality*. Activities related to *manufacturing* tend to have the lowest inequality, as opposed to *finance* and *natural resource extraction*, which show the highest inequality indices. These results, that point to specific industries, are a stepping stone towards a more complete characterization of inequality for policymakers, because it goes beyond analysis at the level of a whole country, and recognizes the role of specific economic activities in generating income inequality.

We found evidence that industrial composition matters to explain inequality in a city. This perspective should therefore be included in further research and policymaking analysis, adding also a measure of *between* inequality for industries. One can start to think of how to foster less unequal industries without sacrificing productivity. This results opens the door to some novel questions. For example, we know that industries differ in the products they produce and the skills they require and cities differ in their size and economic diversity. Consequently, what is the effect of this interaction for income inequality? What is the relationship of the *within* and *between* inequality of industries with other traditional factors such as education and gender?

Future work will also consider informal workers, which is important since inequality is a feature of most economic activities, not just the formal. This dimension needs more analysis, since part of the agenda for a more inclusive growth is understanding the determinants of informality.

Studies that can probe which mechanisms make these industries more, or less, unequal can produce several insights for labor policies that address specific problems of income inequality. Moreover, research that analyzes inequality within a city *and* an industry will provide understanding about an economic activity that is specific to a place. We believe that this point of view, emphasizing the differences between economic activities and places, once developed, will be of great use to practitioners.

References

Berg, A. and J. D. Ostry (2011). Inequality and Unsustainable Growth: Two Sides on the Same Coin? IMF Staff discussion Note 11/08 (Washington: Inrternational Monetary Fund).

Bonilla, L. (2008). Diferencias regionales en la distribución del ingreso en Colombia. *Documentos de trabajo Sobre Economía Regional , No 108*. Banco de la República, Centro de Estudios Económicos Regionales (CEER), Cartagena.

Bonilla, L. (2009). Causas de las diferencias regionales en la distribución del ingreso en colombia, un ejercicio de micro-descomposición. *Documentos de trabajo Sobre Economía Regional , No 111*. Banco de la República, Centro de Estudios Económicos Regionales (CEER), Cartagena.

Credit Suisse Research Institute (2014). Global Wealth Databook 2014. Credit Suisse, Zurich.

Duranton, G. (2013). Delineating metropolitan areas: Measuring spatial labour market networks through commuting patters. Mimeographed, Wharton School, University of Pennsylvania.

Esquivel, G. and, C. G. (2011). The dynamics of income inequality in mexico since nafta. *Economía 12*(1), 155–188.

Galbraith, J. and J. Hale (2014). The Evolution of Economic Inequality in the United States, 1969-2012: Evidence from Data of Inter-industrial Earnings and Inter-regional Incomes. *World Economic Review 3*, 1–19.

Garza, N. (2008). La distribución del ingreso y las economías del caribe colombiano. *Economía, gestión y desarrollo 6*, 245–271.

Marshall, A. (2010). Desigualdad Salarial en la Industria Argentina:Discusió de las Tendencias en 2003-2008. XLV Reunión anual de la Asociación Argentina de Economí Política, Buenos Aires.

Moore, W. (2009). Income Inequality and Industrial Composition. *Public Administration Quarterly 33*(4), 552–581.

OECD (2015). In It Together, Why Less Inequality Benefits All. OECD Publishing, Paris. URL: `http://dx.doi.org/10.1787/9789264235120-en`.

Piketty, T. (2014). Capital in the 21st century. *Cambridge: Harvard University*.

Pineda, J. and L. Acosta (2009). Distribución del ingreso, género y mercado de trabajo en colombia. Mercado Laboral Colombiano. Cuatro Estudios Comparativos. Organización Internacional del Trabajo, OIT, Bogotá, pp. 13-57.

SEDLAC. Socio-Economic Database for Latin America and the Caribbean (CEDLAS and the World Bank).

World Development Indicators. The world bank.