the review of income and wealth

Review of Income and Wealth
Series 66, Number 4, December 2020
DOI: 10.1111/roiw.12425

INEQUALITY IN PRE-INCOME SURVEY TIMES: A METHODOLOGICAL PROPOSAL

BY GUILLERMO LEZAMA*

Facultad de Ciencias Sociales, Universidad de la República, Uruguay

AND

HENRY WILLEBALD

Facultad de Ciencias Económicas y de Administración, Universidad de la República, Uruguay

We propose different alternatives of inequality estimation for economies with a big agricultural sector where land is a decisive factor in income generation and where we do not have enough information about personal earnings. To this end, we use the Uruguayan case to test our methodology. We propose six analytical exercises where Gini indexes are calculated, and as reference we choose the estimation that better adjusts to some theoretical and empirical conditions. Finally, we check the historical accuracy of the series by looking at income distribution explicative variables and the shape of the Inequality possibility frontier. Our results are consistent with the economic and social events of the period (1870–1912) and with previous estimates which reveal worsening trends in income distribution. However, our annual data allow capturing the dynamics of the process where breaks in the series are observed and improvements and declines alternate in the evolution of income distribution.

JEL Codes: N36, O15, D31

Keywords: first globalization, Gini, historical statistics, income inequality, Uruguay

1. INTRODUCTION

Inequality constitutes one of the most frequently discussed topics in social sciences (e.g. Lindert and Williamson, 1982; Persson and Tabellini, 1994; Barro, 2000) and, particularly in economic history and economic development, the debate on the measurement and interpretation of the long run evolution has attracted considerable attention (Deininger and Squire, 1996; Milanovic, 2007).

Part of this debate is fueled by the different measurement concepts where inequality has been defined as—population-weighted—“inter-country inequality” of per capita incomes or as a combination of between- and within-country inequality. In comparative terms, it is possible to identify three concepts about inequality

Note: We have received valuable comments from the session participants in the V Congreso Latinoamericano de Historia Económica, Sao Paulo, Brazil, and we are especially grateful to suggestions of Javier Rodríguez Weber and Pablo Astorga. We also thank to Luis Bértola for sharing his database of personal incomes and number of agricultural workers in 1909–1912 with us and to Leonel Muinelo-Gallo and Carolina Román for their attentive reading. Finally, we thank an anonymous referee and the editor for their suggestions and corrections. All remaining errors are our own.

*Correspondence to: Guillermo Lezama, Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República, Constituyente 1502, Montevideo, 11200, Uruguay (guillermo.lezama@cienciassociales.edu.uy).

(Milanovic, 2012). The first concept is focused on inequality between nations of the world. It is an inequality statistic calculated across per capita GDPs or mean incomes obtained from household surveys of all countries in the world as a proxy of "international inequality". In a second concept, we can correct this measure considering population of each country to obtain a measure of weighted international inequality. Finally, global inequality, which is the most important concept for those interested in the world as composed of individuals (not nations); i.e. each person, regardless of their country, enters in the calculation with their actual income. The most recent article where the changing shape of global inequality in the long run (1820-2000) is studied belongs to Van Zanden *et al.* (2014). In this work—in the tradition of Bourguignon and Morrisson (2002)—the authors apply the main statistical tools for estimating inequality in economic history depending on data availability and periods with the objective of obtaining a consistent dataset of global inequality. This involves the following approaches:

1. Direct estimates of Gini coefficients of income inequality for the post-World War II (WWII) period, when household budget surveys are periodically available, together with efforts to harmonize data basically following the procedure developed by François and Rojas-Romagosa (2005).
2. A large number of estimates of Gini coefficients of income distribution before 1945 are available and the authors converted other measures of income inequality—in particular the numerous estimates of the share of the highest 1 or 5 percent in total income that are available—into comparable Gini coefficients, making use of the assumption that income has a log-normal distribution (Soltow, 1998).
3. In addition, it is possible to apply the idea developed by Williamson (2000a, 2000b, 2002) and followers, and tested by Prados de La Escosura (2008): changes in income inequality in developing countries may be approached by the ratio between real wages and real per capita GDP.
4. Finally, it is possible to assume a relationship between the distribution of heights (a measure of the "biological standard of living") (Steckel, 1995) and income distribution (Baten, 2000, and followers). Such a link can be demonstrated for a set of countries and be used to obtain new data.

Our contribution corresponds to the second and third analytical fields but it differs in two ways from them. Van Zanden *et al.* (2014) test Williamson's ideas for a set of (large) countries, and use this exercise to find the relationship between unskilled wages and per capita GDP in order to extrapolate or intrapolate Gini coefficients for a sample of countries for which the authors do not have direct estimates. By contrast, we consider only one case and a different Williamson Index: land rent /unskilled wage (r/w).

Considering only one index in this field is not new. Prados de la Escosura (2005, 2007) propose Gini coefficients projected backwards with inequality indices constructed as the ratio between unskilled wage indices and GDP per worker. He obtains Gini coefficients for 10-year periods and four Latin American countries: Argentina, Brazil, Chile and Uruguay, from the second half of the 19th century onwards.

Our interest is to prove the convenience of using an alternative index according to the local economic and social conditions. In those cases where a component of the economy is a big agricultural sector where land is a critical component of total wealth and a decisive factor in income generation, and where the landowning class is a minority (elite), the use of *r/w* ratio is preferable. This index considers trends in the ratio between farm rents per unit of hectare and the wage rate of the lower qualification tasks, which can be understood as a measure of how many days an employee has to work to pay the rent for one unit of land. We do not use this index directly although it is considered to be representative of income distribution evolution in several places of the world periphery (as Williamson, 1999, 2000a, 2000b; Álvarez-Nogal and Prados de la Escosura, 2013; and Milanovic, 2016, demonstrate). Instead, and inspired by this idea, we use trends in individual wages and land rents per earner (not per hectare) separately. In addition, our proposal considers annual estimations of Gini coefficients which allow approaching the dynamics of the process and evaluating the accuracy of our estimation. However, estimating Gini coefficients annually, for the second half of the 19th century and taking as reference the r/w ratio are not new.

Bértola (2005) proposes an estimation of Gini coefficients year by year for Uruguay in the period 1880–1908 as a "working hypothesis" (or "provisional synthesis"). His estimation is based on a logarithmic transformation of the series land prices/wages (an index very close to *r/w* ratio) as an exercise that tries to reflect the sign of the trend, not the proportional changes. Uruguay is an excellent case to test our methodology because: (1) the characteristics of the economy are suitable to prefer *r/w* ratio; (2) it is a relatively homogenous territory in productive and distributive terms; (3) we have previous estimates to contrast and evaluate our results; (4) information availability is higher than other peripheral economies and it allows working with additional data to improve our estimation.

As a settler economy (Álvarez *et al.*, 2007; Lloyd and Metzer, 2013) Uruguay presents proper characteristics to apply *r/w* index. Its economic and social development articulated dynamic relations between waves of immigration, marginalization of native people, European capital inflows, land abundance, free labour, socially-useful political institutions and neo-European cultures. By the late 19th century it was well integrated into the global economy thanks to the production of raw materials and food derived from land exploitation and institutional arrangements that guaranteed landownership.

Previous studies (Martínez-Galarraga *et al.*, 2019) have demonstrated that Uruguay presents, in the long run, differences in the level of territorial development. However, these differences were lower than other settler economies with larger territories and bigger discrepancies in terms of natural endowments, climatic conditions and incorporation of overseas production factors (for instance, Argentina, Australia or Canada). Therefore, we can obviate regional differences in our estimates and consider that our index is representative of the whole economy.

Two estimates are available for the period—Bértola (2005) and Prados de la Escosura (2007)—and both confirm the Stopler-Samuelson interpretation. Williamson (1999) explores the consequences for inequality of the First Globalization (1870–1914). On the basis of the wage-land rental ratio, he shows an increase in inequality within Uruguay in accordance with that theoretical framework. As natural resources were the abundant production factor in the majority of Latin American countries, they were more intensively used in the production of exportable commodities. As a result, returns to land grew relatively to those of labor. Since ownership of natural resources is more concentrated than that of labor, income distribution tended to be skewed towards landowners and inequality rose over the decades prior to World War I (WWI).

Finally, our methodology needs additional data apart from *r*/*w* ratio to obtain results: total unskilled agrarian workers (people who earn wages); total landowners (people who earn rental income), economically active population (as a good proxy to people who earn incomes) and an estimation of total GDP (which is the total income to distribute). We can use these variables in the Uruguayan case because estimates are available.

In Section 2, we show how the Gini Index, assumed as a probability function, can be linked with income shares, and we propose indexes with six different exercises. We also propose to check their robustness with several methods. In Section 3, we present our data sources and some assumptions we make in order to use it as we do not have annual data for all the periods. In Section 4, we show how we calculate the Gini Index with those exercises and present the results. For all exercises, we compare our estimates with previous ones, and we also check if they are reliable according to other historical evidence. In Section 5, with our final estimates, we check the historical accuracy of the index by looking at income distribution explicative variables. We also use the inequality possibility frontier to check the robustness of our results. Finally, we conclude and propose a research agenda (Section 6). Our results indicate that in the 1870–1912 period, Uruguay shows an increasing long term income inequality. Our estimates also indicate that Uruguayan income distribution shape is more like a Pareto function distribution than a lognormal one.

2. Our Methodological Proposal

2.1. *Components of the Estimation*

One of the aims of this work is to obtain an inequality indicator (Gini Index) for Uruguay during the First Globalization (1870-1913) using the scarce data that we have: agrarian wage rates, land rent rates, shares of agrarian workers and landowners (owning more than 100 hectares) on the total economically active population (EAP) and the entire GDP (in current prices) as representative of total income.

Assuming that agrarian workers and landowners belong to the poorest and the richest segments of the EAP, respectively, we know the upper and lower tails of income distribution.[1] To calculate the Gini coefficients we need the whole income distribution and we propose assumptions and analytical exercises to obtain inequality indicators with information about the tails of the distribution only.

[1]It may be discussable that landowners were part of the EAP. We consider it in this way because of two reasons: (1) we need representing the whole income earners; (2) In the River Plate the landownership absentee was an extended phenomenon but we do not know how many land proprietors corresponded to these condition for all the period; many landowners occupied the land and produced actively leaving that condition, especially, to owners of large estate (*latifundistas*).

This characterization is incomplete because perhaps landowners were not the richest "social class" and the agrarian workers were not the poorest one in the society at that time. In other words, as there could be people who were richer than landowners and others who were poorer than agrarian workers, our frame is imperfect. In addition, it would be possible to find rich people among landowners that obtain income from other sources—typically, from international trade business—and poor people with occupations different from "*peones*". We will use this argument to check the robustness of our results.

In our estimates, we consider six components and, in general, we need to know or assume five of them as "true" (our degrees of freedom) leaving the sixth component as one of the results (the second result will always be the Gini Index).

1. Number of poor people (agrarian workers: L_W).
2. Per capita income corresponding to poor people (agrarian wage rate: w).
3. Number of rich people (landowners: L_R).
4. Per capita income corresponding to rich people (land rent rate: r).
5. Total income (GDP).
6. The shape of distribution.

Two types of per capita income distributions are usually applied in the literature about historical inequality—Lognormal and Pareto—and we test both functions in our exercises.[2] Then, we alternate assumptions related to incomes of poor and rich people and GDP according to the possibilities offered by the methodology.

2.2. *Total Income and Income Earners*

Functional income distribution is a depiction of how total income (Y) is distributed among the different groups involved in production. As a result, it shows how incomes earned by the owners of the various production factors (labor, land and capital) are shared out in terms of remunerations (or wages), land rents and profits (dividends or interests). Functional distribution of income is a good proxy for interpersonal income distribution in early stages of development because each component resulted relatively homogenous within the corresponding group. Then, we can define total income as the sum of wages (W), land rents (R) and benefits (B).

$$Y = W + R + B \tag{1}$$

Wages include earnings of unskilled and skilled rural and urban workers (four types of wages). We assume that in a settler economy as Uruguay, the poorest social class is composed by unskilled rural workers and the richest one is composed of landowners, who earn, respectively, the lowest wage (w) and land rents (r). Therefore, we can rewrite equation (1) including both social classes corresponding to the tails of the income distribution and the rest of earnings as a set of income.

[2]Another possibility would be using more flexible distributions, but these rely on the availability of more data because we should estimate more than one parameter. That is the reason why we relied on these distributions since we only need one parameter in order to estimate the Gini index.

$$Y = W_{UnskRural} + R + Others \tag{2}$$

We divide by total income,

$$1 = a_w + a_r + a_{others} \tag{3}$$

In the primary distribution of income, each production factor is rewarded according to the income generation that results directly from the production process, and the distribution of it over the production factors. These earners of incomes (L)—that we identify with the EAP–can be divided in terms of workers (L_w), land renters or landowners (L_R) and the rest. Therefore,

$$L = L_W + L_R + L_{others} \tag{4}$$

We divide by the total income earners to obtain the shares of each type,

$$1 = s_w + s_r + s_{others} \tag{5}$$

Previously, we argued that r/w is a main concept to explain the evolution of income distribution in settler economies. We express w and r as follows,

$$w = \frac{W}{L_W} \tag{6}$$

$$r = \frac{R}{L_R} \tag{7}$$

Then,

$$\frac{r}{w} = \frac{\frac{R}{L_R}}{\frac{W}{L_W}} = \frac{R}{L_R} \cdot \frac{L_W}{W} = \frac{R}{L_R} \cdot \frac{L_W}{W} \cdot \frac{Y}{Y} \cdot \frac{L}{L} = a_r \cdot \frac{1}{a_w} \cdot s_w \cdot \frac{1}{s_r}$$

$$\frac{r}{w} = \frac{a_r}{a_w} \cdot \frac{s_w}{s_r} \tag{8}$$

In other words, if we assume that unskilled rural workers and landowners (land renters) are, respectively, the poorest and the richest classes, we can follow the r/w ratio to approach the evolution of relative incomes (weighted by active population). For the next steps in the presentation, it is useful to clear up the relationship among accumulated incomes of the tails of the distribution.

$$\frac{a_w}{a_r} = \frac{w}{r} \cdot \frac{s_w}{s_r} \tag{9}$$

2.3. *Shapes of Income Distribution*

Pareto Income Distribution

According to Moothathu (1985), the Gini index (G) can be calculated as,

$$G=\frac{1}{(2k-1)} \tag{10}$$

where k refers to the coefficient that defines the lower tail of the distribution (i.e. the earnings of the lower class):

$$a_w=1-(1-s_w)^{(k-1)/k} \tag{11}$$

But k refers, also, to the coefficient that defines the upper tail of the distribution (i.e. the earnings of the higher class):

$$a_r=1-\left(1-(s_r)^{(k-1)/k}\right)=(s_r)^{(k-1)/k} \tag{12}$$

Lognormal Income Distribution

According to Lopez and Servén (2006), the Gini index (G) is given by,

$$G=\Phi\left(\frac{\sigma}{\sqrt{2}}\right).2-1 \tag{13}$$

where,

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}.e^{\frac{-t^2}{2}}dt \tag{14}$$

where σ refers to the coefficient that defines the lower tail of the distribution:

$$a_w=\Phi\left(\Phi^{-1}(s_w)-\sigma\right) \tag{15}$$

and, also, it refers to the coefficient that defines the upper tail of the distribution:

$$a_r=1-\left(\Phi\left(\Phi^{-1}(1-s_r\right)-\sigma\right) \tag{16}$$

2.4. *Empirical Exercises*

We consider, for each type of distribution, only those cases where we assume values for four of the remaining five components of the estimation; it is possible to propose exercises assuming values for three cases but we need additional assumptions. In Table 1, we tick those variables whose values we assume (or know), and the rest will be the results of the estimation.

In brief, our exercises refer to calculate the Gini Index knowing (1) the inferior and superior tails of the distribution but not the GDP or total income (exercises 1 and 2); (2) the inferior tail of the distribution and the GDP, but not the superior tail (exercises 3 and 4); and (3) the superior tail of the distribution and the GDP, but not the inferior tail (exercises 5 and 6).

3. Data

We make use of novel available data about inequality in Uruguay for the First Globalization. Willebald (2011, 2015) offers data about functional income distribution in agriculture for five time benchmarks—1874, 1883, 1893, 1903 and 1912—which constitute our main reference. Data refer to the structure of the total income according to the distribution among renters (land rents), workers (wages) and entrepreneurs (benefits). We apply the resultant structures to the agriculture GDP (3-year averages) reported in Bonino *et al.* (2012) to obtain the corresponding incomes in current prices.

Original wage series includes unskilled wage-earners (*"peones"*), foremen (*"capataces"*) and servants (*"sirvientes"*) as the total agrarian workers, but we only consider the former as the poorest segment of society of that time. We know the corresponding personal incomes and number of workers of each category in 1909–1912 (Bértola, 2005) but only the total incomes and the total workers for the rest of the time-benchmarks. Then, we use the income (69 per cent) and number (84 per cent) of shares of that period referred to *"peones"* to estimate the corresponding total workers in the previous benchmarks. We obtain annual data by interpolation and the wage rates, year by year, dividing both concepts. An alternative is to use a wage index to scale and join the wage rates of the different benchmarks. However, the available index (Bértola *et al.*, 1999) presents an excessive urban profile (Willebald, 2011, 2015) that does not fit appropriately with our agrarian data. In particular, they show inverse trends between benchmarks and this is why we disregard the wage index.

Original land rent series came from Willebald (2015) considering five time benchmarks (1874, 1883, 1893, 1903 and 1912). The number of renters (that we identify as landowners) came from Bértola (2005), corresponding to plots larger than 100 hectares and we consider the average 1909-1912 as reference. These data distinguish between landowners by type of production—livestock and crops—and

TABLE 1
Assumptions for Empirical Exercises

	Distribution	L_W	W	L_R	R	Y
Exercise 1	Pareto	√	√	√	√	
Exercise 2	Lognormal	√	√	√	√	
Exercise 3	Pareto	√	√	√		√
Exercise 4	Lognormal	√	√	√		√
Exercise 5	Pareto	√		√	√	√
Exercise 6	Lognormal	√		√	√	√

we obtain the total renters for the rest of the benchmarks according to the movement of the area (retropolation) corresponding to each activity (data from Willebald, 2011). The ratio between both series renders land rent rates[3] for each benchmark and we obtain annual data splicing and rescaling these temporal points with a land rent index (Bértola, *et al.*, 1999).[4]

4. Results

4.1. *Assumption 1: We Know the Total Incomes of "Poor" and "Rich" People*

As an illustration of our methodological approach, we present our outcomes compared with two antecedents—Prados de la Escosura, 2007, for benchmark years; Bértola, 2005, for annual data and evolution—and we control our estimates according to these two previous results.

On the one hand, Bértola (2005) presents estimates of Gini coefficients for the 1908–1968 period made with the methodology of social tables (in the tradition of Lindert and Williamson, 1982, and followers). This offers a set of high quality estimates for historical analysis that we use as reference. For instance, according to Bértola (2005), the Gini coefficient corresponding to 1908–1910 is 0.37, and our estimate is 0.35 and 0.34 (for Pareto and Lognormal distribution respectively). The proximity of the results reinforces the fact that our method renders interesting insights.

On the other hand, in fact, when we "distribute" incomes among earners we should reproduce the total income or GDP; we contrast the actual estimates of GDP for the period and our implicit total GDP to evaluate the accuracy of our results (see below).

Pareto Income Distribution

Considering equation (9) and the results of equations (11) and (12), we can operate as follows,

$$\frac{a_w}{a_r} = \frac{1-(1-s_w)^{(k-1)/k}}{(s_r)^{(k-1)/k}} \tag{17}$$

$$\frac{a_w}{a_r} = \frac{1}{(s_r)^{(k-1)/k}} - \frac{(1-s_w)^{(k-1)/k}}{(s_r)^{(k-1)/k}} \tag{18}$$

$$\frac{a_w}{a_r} = \left(\frac{1}{s_r}\right)^{(k-1)/k} - \left(\frac{1-s_w}{s_r}\right)^{(k-1)/k} \tag{19}$$

[3]This ratio is a proxy of the rent per hectare because the total area destined to agricultural production was relatively stable in the period.

[4]We use moving 5-year averages of the indexes to reduce steep changes in the original series.

According to (9)

(20)
$$\frac{w}{r}\cdot\frac{s_w}{s_r}=\left(\frac{1}{s_r}\right)^{(k-1)/k}-\left(\frac{1-s_w}{s_r}\right)^{(k-1)/k}$$

(21)
$$0=\left(\frac{1}{s_r}\right)^{(k-1)/k}-\left(\frac{1-s_w}{s_r}\right)^{(k-1)/k}-\frac{w}{r}\cdot\frac{s_w}{s_r}$$

We obtain k by the Newton-Raphson method and obtain the Gini Index G (equation 10) year by year.

Lognormal Income Distribution

Combining expressions of equations (15) and (16),

(22)
$$\frac{a_r}{a_w}=\frac{1-\left(\Phi\left(\Phi^{-1}(1-s_r\right)-\sigma\right)}{\Phi\left(\Phi^{-1}(s_w\right)-\sigma)}$$

(23)
$$\frac{a_r}{a_w}=\frac{1}{\Phi\left(\Phi^{-1}(s_w\right)-\sigma)}-\frac{\left(\Phi\left(\Phi^{-1}(1-s_r\right)-\sigma\right)}{\Phi\left(\Phi^{-1}(s_w\right)-\sigma)}$$

As we know s_w and s_r we can calculate the corresponding values in function of Φ^{-1}. If we name,

$$\Phi^{-1}\left(s_w\right)=l$$

$$\Phi^{-1}\left(1-s_r\right)=j$$

Then,

(24)
$$\frac{a_r}{a_w}=\frac{1}{\Phi(l-\sigma)}-\frac{\Phi(j-\sigma)}{\Phi(l-\sigma)}$$

To calculate σ we follow Vázquez-Leal *et al.* (2012) where the homotopy perturbation method is applied to approximate the cumulative distribution function of a standard normal random variable.[5] The analytical expression is the following,

(25)
$$\Phi(x)\approx\left[e^{\left(\frac{-358x}{23}+111arctan\left(\frac{37x}{294}\right)\right)}+1\right]^{-1},\ -\infty\leq x\leq+\infty,$$

Then, we can rewrite the expression (24),

[5]This approximation is very precise for $x > -3$

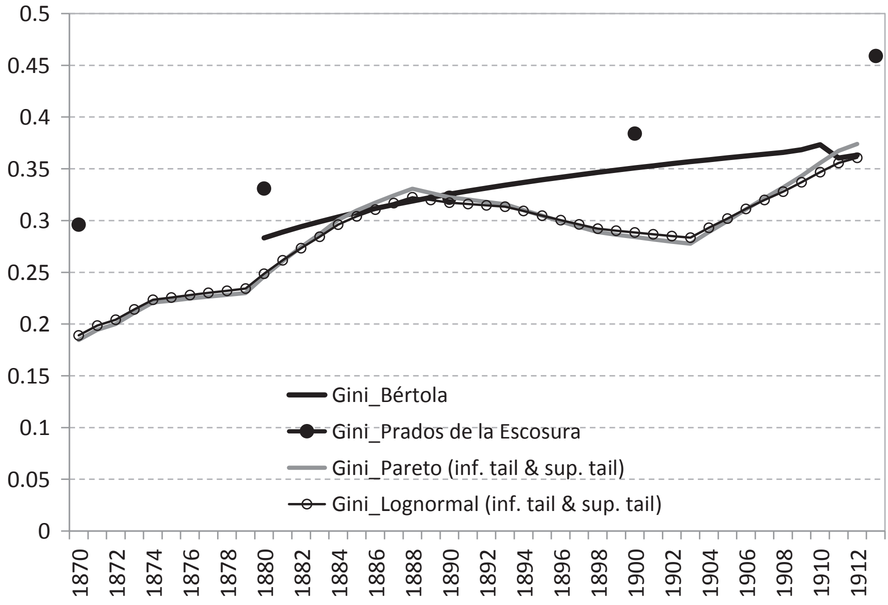

Figure 1. Gini coefficients: previous results and our initial estimates
Source: Bértola (2005), Prados de la Escosura (2007) and own estimates.

$$(26)\quad \frac{a_r}{a_w} \approx \frac{1}{\left[e^{\left(\frac{-358(l-\sigma)}{23}+111arctan\left(\frac{37(l-\sigma)}{294}\right)\right)}+1\right]^{-1}} - \frac{\left[e^{\left(\frac{-358(j-\sigma)}{23}+111arctan\left(\frac{37(j-\sigma)}{294}\right)\right)}+1\right]^{-1}}{\left[e^{\left(\frac{-358(l-\sigma)}{23}+111arctan\left(\frac{37(l-\sigma)}{294}\right)\right)}+1\right]^{-1}}$$

We obtain σ and, then, we calculate G in equation (13).

Now, we present our results. In Figure 1 we represent the evolution of income distribution in Uruguay during the First Globalization according to Bértola (2005), Prados de la Escosura (2007)[6] and our results, which include estimates corresponding to assumption referred to the inferior and superior tails of the income distribution. We assume two shapes for this distribution: Pareto and Lognormal.

First, we obtain the same increasing long term trend. In other words, at the moment, the evidence about an increasing income inequality in Uruguay during the First Globalization seems unquestionable (as in other settler economies). Second, we obtain, with the exception of the end of the 1880s, lower levels of Gini coefficients and, specially, from the second half of the 1890s to the beginning of the 20th century. Our assumptions can be conservative in the sense that they offer a "floor" of inequality; we will use this consideration in next estimates. Third, we find a break in that rising trajectory that is also insinuated in Prados de la Escosura's data. According to our methodology, the GDP is another result of our estimation. In Figure 2 we present the current series of Uruguay's GDP for the period and the corresponding estimates for both types of income distribution.

[6]Table A1 in online Appendix shows, in column 1 and column 2, previous estimates from Bértola (2005) and Prados de la Escosura (2007).

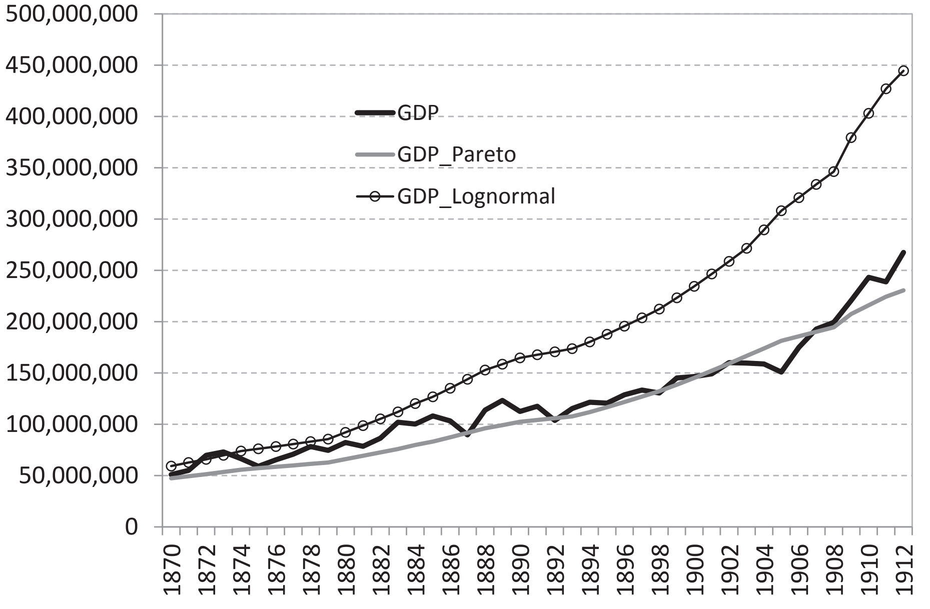

Figure 2. GDP and implicit GDP in our estimates
Source: Bonino *et al.* (2012) (based on Bértola, 1998; Bertino & Tajam, 1999; and Official National Accounts, BCU) and own estimates.

Evidently, in terms of the Gini index, assuming a Pareto distribution and a Lognormal distribution renders similar results, but the outcomes in terms of GDP are significantly different. The average GDP corresponding to a Pareto income distribution is 94 percent of the actual GDP and that corresponding to Lognormal income distribution exceeds the actual GDP by 55 per cent. This different result in terms of GDP is not a simple coincidence. We can state that if we have similar Gini indexes, as is our case for all years (Figure 1), GDP will be higher in the Lognormal based estimates of GDP than in the Pareto based estimates. This is consequence of the shapes of distributions; in Online Appendix 1 we propose some simulations for supporting our argument.

4.2. *Assumption 2: We Know GDP and Total Incomes of Poor People*

If we consider as true the total GDP and we know the inferior tail of the distribution, we will obtain as a result the superior tail.

Pareto Income Distribution

Assuming a Pareto distribution and considering again equation (11), we operate,

(27) $$(1-s_w)^{(k-1)/k} = 1-a_w$$

(28) $$ln(1-s_w)^{(k-1)/k} = ln\left(1-a_w\right)$$

$$((k-1)/k)\,.ln\left(1-s_w\right)=ln\left(1-a_w\right) \tag{29}$$

$$\frac{k-1}{k}=\frac{ln\left(1-a_w\right)}{ln\left(1-s_w\right)} \tag{30}$$

$$1-\frac{1}{k}=\frac{ln\left(1-a_w\right)}{ln\left(1-s_w\right)} \tag{31}$$

$$\frac{1}{k}=1-\frac{ln\left(1-a_w\right)}{ln\left(1-s_w\right)} \tag{32}$$

$$k=\frac{1}{1-\frac{ln\left(1-a_w\right)}{ln\left(1-s_w\right)}} \tag{33}$$

With k, we obtain the corresponding Gini index—equation (10)—and with k and GDP we obtain r—equation (12) which allows determining the superior tail of the distribution.

Lognormal Income Distribution

Assuming a Lognormal distribution and considering again equation (15),

$$a_w=\Phi\left(\Phi^{-1}(s_w)-\sigma\right)=\Phi\left(l-\sigma\right)$$

$$\Phi^{-1}(a_w)=l-\sigma \tag{34}$$

$$\sigma=l-\Phi^{-1}(a_w) \tag{35}$$

With σ, we obtain the corresponding Gini index—equation (13)—and with k and GDP we obtain r—equation (16)—which allows determining the superior tail of the distribution.

Again, we obtain the same increasing long term trend but now the trajectory is not so smooth as before (Figure 3). On the contrary, we can distinguish different income distribution evolutions where the changes that we identified previously are now more pronounced. Both types of distribution follow a similar trajectory but the Lognormal distribution renders a lower Gini index than that corresponding to Pareto distribution. These low levels seem barely believable for some periods—for instance, they are between 0.15 and 0.20 in the 1890s and close to 0.10 in the beginning of the 20th century—and, compared with our reference for 1908–1910, the value is excessively small (0.21 vs. 0.37 of Bértola, 2005). However, the Pareto distribution renders a value closer to our reference (0.37) which reinforces the feasibility of this approach.

4.3. *Assumption 3: We Know GDP and Total Incomes of Rich People*

If we assume that the total GDP is true and we know the superior tail of the distribution, we will obtain as result the inferior tail.

Pareto Income Distribution

Assuming a Pareto distribution and considering equation (12), we operate,

$$\ln a_r = ln(s_r)^{(k-1)/k} \tag{36}$$

$$\ln a_r = \frac{k-1}{k} ln(s_r) \tag{37}$$

$$\frac{\ln a_r}{\ln s_r} = 1 - \frac{1}{k} \tag{38}$$

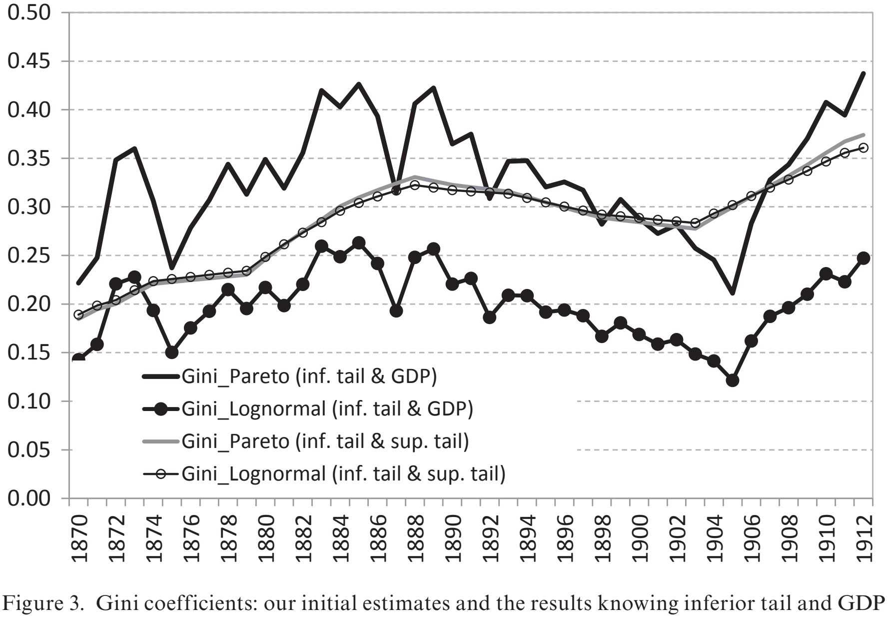

Figure 3. Gini coefficients: our initial estimates and the results knowing inferior tail and GDP
Source: own estimates.

$$1 - \frac{\ln a_r}{\ln s_r} = \frac{1}{k} \tag{39}$$

$$k = \frac{1}{1 - \frac{\ln a_r}{\ln s_r}} \tag{40}$$

With k, we obtain the corresponding Gini index—equation (10)—and with k and GDP we obtain r—equation (12)—which allows determining the superior tail of the distribution.

Lognormal Income Distribution

Assuming a Lognormal distribution and considering equation (16),

$$a_r = 1 - \left(\Phi\left(\Phi^{-1}(1 - s_w\right) - \sigma\right) = 1 - \Phi\,(j - \sigma) \tag{41}$$

$$1 - a_r = \Phi\,(j - \sigma) \tag{42}$$

$$\Phi^{-1}\left(1 - a_r\right) = (j - \sigma) \tag{43}$$

$$\sigma = -\Phi^{-1}\left(1 - a_r\right) + j \tag{44}$$

With σ, we obtain the corresponding Gini index—equation (13)—and with k and GDP we obtain r—equation (15)—which allows determining the superior tail of the distribution.

We obtain an increasing long term trend and less smooth than before. Both types of distributions follow a similar trajectory but the Lognormal distribution renders a higher Gini index than that corresponding to Pareto distribution. These high levels render a Gini index of 0.49 in 1908–1910 that seems excessively big (although close to Prados de la Escosura's figure). On the other hand, the Pareto distribution shows a value (0.33) closer to our reference (0.37) which reinforces our idea that this kind of distribution would describe the inequality at that time more accurately (see Figure 4).

4.4. *Reliability of Our Estimates*

According to the previous exercises, we can accept that the shape of the income distribution is Pareto but we have three estimations. Which evolution and levels of the Gini index are the most credible ones? To answer this question, we analyze how feasible our Gini indexes are.

We discard Exercise 1 because we assume that the available historical series of GDP (Bonino *et al.*, 2012) is a suitable estimation for our aim.

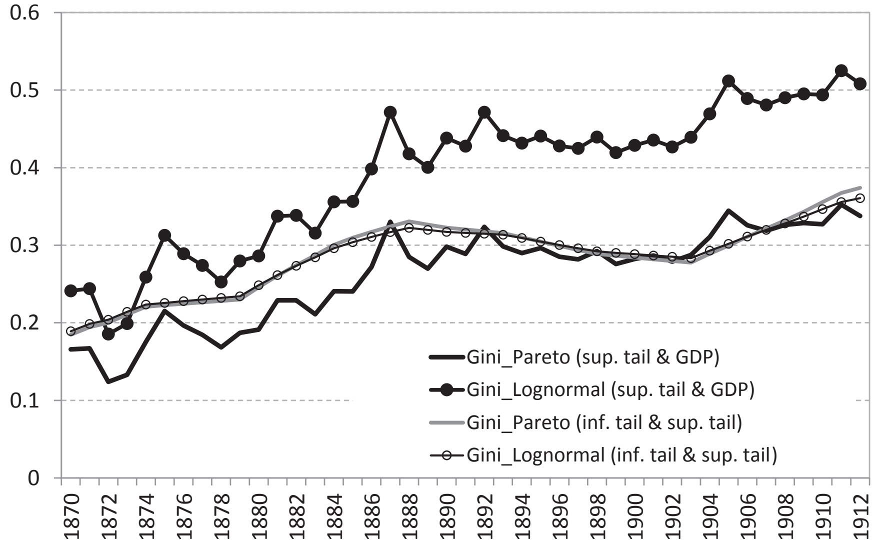

Figure 4. Gini coefficients: our initial estimates and the results knowing superior tail and GDP
Source: own estimates.

In Exercise 3—when we know GDP and inferior tail—we deduce the upper tail of the income distribution (i.e. the income of the upper class). This result should be equal to or higher than the actual value of this modality of income because a discrepancy in the other sense would mean that there are people in the upper class who earn less income than our landowners, which is inconsistent with the initial assumption (as we know the total renters, we would be admitting an implicit, and non-possible, per capita income of the upper class lower than the actual *r*). Analogously, in Exercise 5—where we know the GDP and the superior tail—we obtain a share of the lower tail of the income distribution and this result should not be higher than the actual value because in this case, it would mean that poor people would obtain a higher income than our workers' earnings (*w*), and that is not consistent with our initial assumption.

We construct two "feasibility indexes" by comparing estimated with actual shares of land rents and wages (Figure 5). Estimated incomes lower than actual incomes are not feasible figures and this happens when indexes are less than 1.

Therefore, estimates of the Gini Index with a Pareto income distribution and considering as "true" the inferior tail of the distribution and the GDP (Exercise 3) are feasible for the entire period with the exception of 1887, 1892, 1898, 1901, and 1903–1906 (see continuous line in Figure 5). For these years the same type of distribution is feasible but with the assumption that the superior tail is true (Exercise 5). The final estimates are the combination of both series.

These estimates are close to Bértola's data for the period 1908–1912. Our data is more variable for that period, but it reflects the same direction in annual variations. Therefore, we decided to splice our estimates with Bértola's data from 1911 onwards. To smooth this splicing we considered the average of 1908–1910 as reference. We represent the combination of those series in Figure 6 and we report

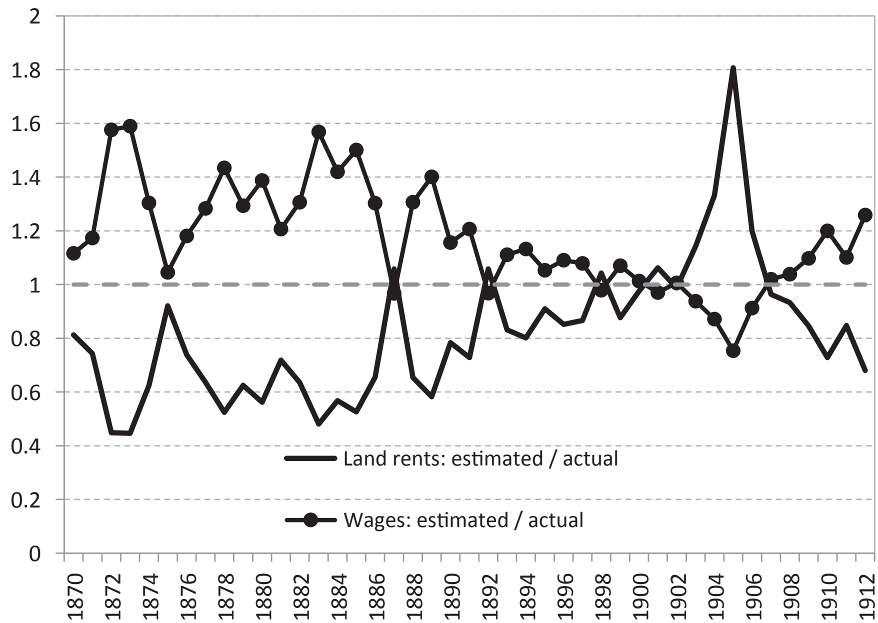

Figure 5. Feasibility indexes: land rents (estimated/actual) and wages (estimated/actual)
Source: own estimates.

these results in the third column of Table A1 in online (with the results before the splicing between parentheses).

Finally, it is important to mention that our estimates show a remarkable correlation with Williamson Index for this period (see Online Appendix 2), but our methodology allows estimating levels of Gini indexes without requiring long-run retropolation of (modern) Gini indexes (calculated from survey data) (as Prados de la Escosura, 2005, does). This issue helps to validate our estimates in its dynamics. We think, therefore, that both methodologies can be excellent complementary tools to deal with historical analysis of inequality.

5. Historical Accuracy of Our Estimates

Statistical feasibility does not assure the historical accuracy of our estimates. Are the evolution and the levels of our index consistent with the historical facts? Or, in other words, were the changes in the trajectory of inequality—within a worsening long-run trend—in accordance with the movements of the expected explicative variables?

As we mentioned previously, the debate in economic history literature about the evolution of income distribution has been based, mainly, on the static neoclassical trade theory developed by Eli Heckscher and Bertil Ohlin in the early years of the 20th century. It is not the aim of this paper to discuss the evolution of inequality in Uruguay during the 19th century—this will be part of the next steps in our research—but to present an alternative estimation methodology of income

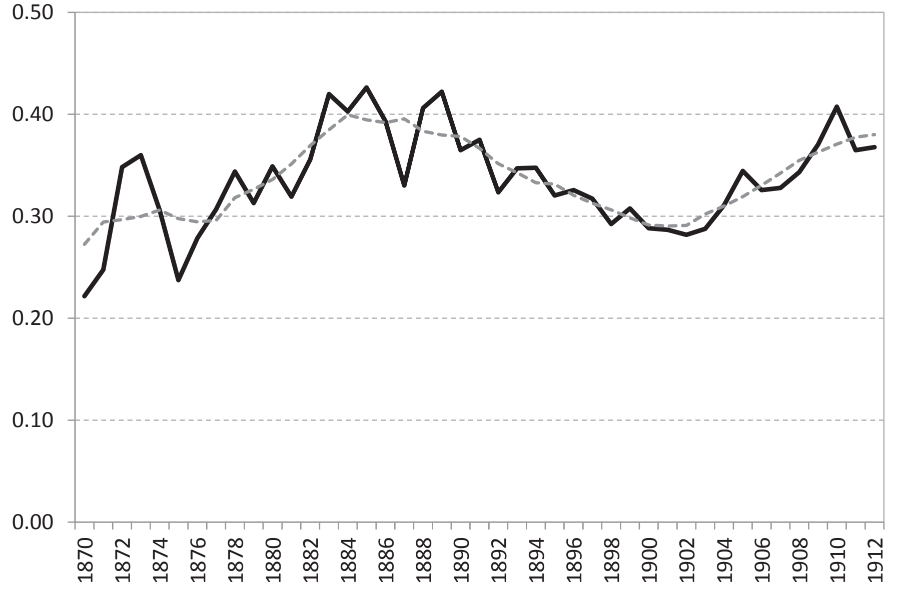

Figure 6. Gini Index of Uruguay during the First Globalization: our proposal (original data and moving 5-year average)
Source: own estimates.

distribution indicators. Therefore, we only review some stylized facts to give credibility to our estimates.

The basic insight of that framework was that trade patterns reflect differences in endowments across economies, and countries export goods embodying those factors of production with which they are well-endowed. Commodity market integration therefore leads to an increase in the demand for abundant (and, consequently, cheap) factors of production, thus increasing their prices, while trade leads to a decline in the demand for scarce (and, consequently, expensive) factors of production which reduce their prices.

The late 19th century was characterized by dramatically declining transport costs, mass migration from the Old World to the New and by large transfers of capital with similar direction. These are the stylized facts of the First Globalization (O'Rourke, 2001). How did each of these separate dimensions of globalization influence income distribution within countries?

Commodity-price convergence—expressed by the ratio of agricultural prices to manufacturing prices (P_a/P_m) or, equivalently for the periphery countries, the ratio of export prices to import prices (P_x/P_{im})–has been associated with relative factor-price convergence (that is, the convergence between *w*/*r* in the Old and the New World). If the relative price of the commodities converges between trading partners, the *w*/*r* should also converge; that is, it should fall in the land-abundant and labor-scarce country (since the export boom raises the relative demand for land), and it should rise in the labor-abundant and land-scarce country (since the export boom raises the relative demand for labor). Where land was held by the favored few and where industrialization had not yet taken hold, the pre-WWI commodity-price convergence implied greater inequality in resource-abundant economies like those in the Southern Cone. It also implied lower inequality in resource-scarce economies like those in Western Europe. The impact of mass migration reinforced this trend.

In the Atlantic economy, real wages and living standards converged from the mid-19th century until WWI. This process was driven by the narrowing of the wage gap between the New and the Old World. In addition, many European countries, particularly the poorer ones, were catching up with the economic leaders in Europe (the industrial countries). Migration affected long-run equilibrium output and wages through changes in aggregate labor supply; it raised wages in countries with high emigration rates and reduced them in countries that received migrations. Capital flows acted as an anti-convergence force (in the sense of the Lucas Paradox) because they moved towards rich countries, rather than poor ones, in pursuit of abundant natural resources, young populations, and the (potential) abundance of human capital (Clemens and Williamson, 2004). Therefore, in contrast with the other factors, capital-deepening in the nonfarm sector of the New World should have drawn labor off the land and raised the wage-rental ratio improving income distribution. Diagram of Table 2 shows a schematic representation of these relations.

We propose some exercises of time consistency between explicative variables and inequality although they are far from being conclusive. A rigorous analysis will be the subject of another paper. The aim is only to find evidence that contributes to give feasibility to our estimates.

First, we evaluate the incidence of the "forces of globalization" measured, as Williamson (2002) proposes, by changes in the relative prices (Figure 7). The relation between the Uruguayan terms of trade (P_x/P_{im}) and the Gini Index are in accordance with the theory. This ratio increases until 1891–1892, which coincides with the last higher peaks in the Gini Index. The posterior decline in inequality coincides with the stability in the index and, then, the evolutions are very similar, confirming the positive relationship we expected. However, this is not true for P_a/P_m.

It is from the 1890s onwards that the ratio between agricultural and manufacturing prices increased, and it just evolved according to the evolution of export-import prices from the beginning of the 20th century. This comparative evolution reveals two facts. On the one hand, it shows that there was a gap between international and domestic formation of prices in spite of the increasing integration of

TABLE 2
Stylized Facts of First Globalization in Land-Abundant Economies and Consequences in Terms of Inequality

If	Explicative Variable		Inequality
⇧	P_a/P_m	⇒	⇧
⇧	P_x/P_{im}	⇒	⇧
⇧	Immigration	⇒	⇧
⇧	Capital inflows	⇒	⇩

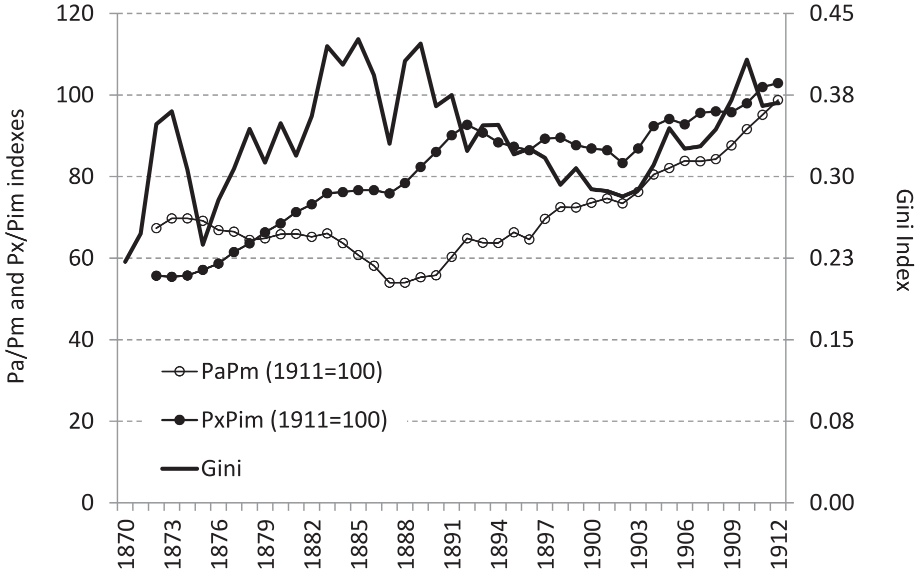

Figure 7. Commodity-price convergence and inequality Pa/Pm, Px/Pim (5-years average) and the Gini Index
Source: Williamson (2000b, 2002) and own estimations.

Uruguay to the world markets. On the other hand, and related with the last point, inequality responded more intensively to international than to domestic conditions.

Second, the rising inequality of the second half of the 1870s and the 1880s coincides with a stage of increasing immigration in Uruguay (Figure 8). The posterior decrease in inequality and the recovery at the end of the period goes hand in hand with similar evolutions of immigration rates. Like other economies of European recent settlement, Uruguay received many immigrants but, contrary to Argentina, it evidenced important departures of aliens. Therefore, we include the net migration rate to represent the net impact of the process. Clearly, the evolutions of both rates are similar but at different levels, and the net rates achieved values close to zero at the beginning of the 20th century.

Third, we do not have complete information about capital inflows. We use as a proxy the British capital exports to Uruguay (annual average in moving 5 years) expressed in per capita terms (Figure 9). We expect a negative relation between this variable and the inequality levels. However, we observe a pattern with positive relations between both variables. This finding is not new. Williamson (2000b) demonstrates that capital-deepening improves inequality for his entire sample of economies,[7] but he obtains the inverse (significant) result for the New World.[8] He rationalizes this result by an appeal to a significant labor-saving technological change in these countries which, in the case of Uruguay, could be associated with the wire-fencing process of the 1870s and 1880s (Millot and Bertino, 1996, p. 61).

[7]His sample is composed of 19 countries.
[8]Argentina, Australia, Canada, Uruguay, and USA.

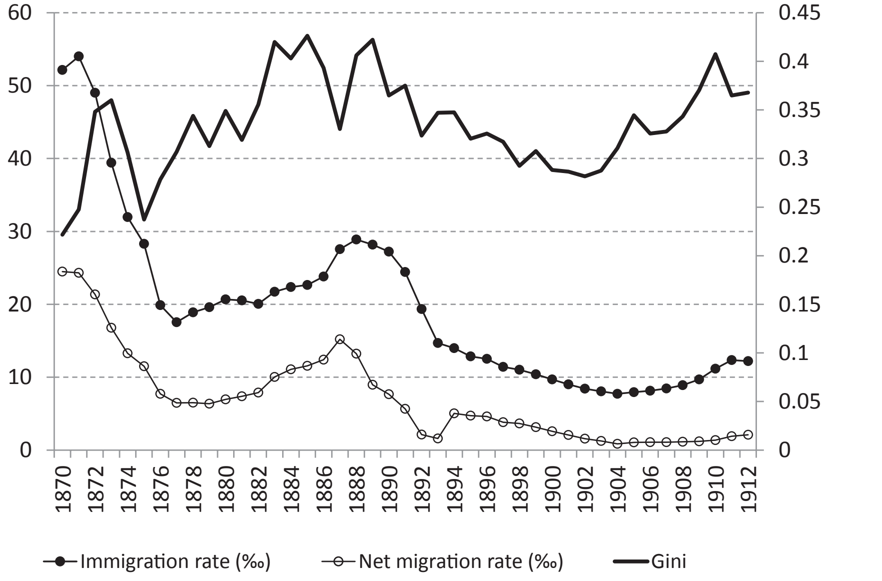

Figure 8. Migration and inequality immigration and net migration rates (‰) and the Gini Index
Source: Mitchell (1993) and own estimates.

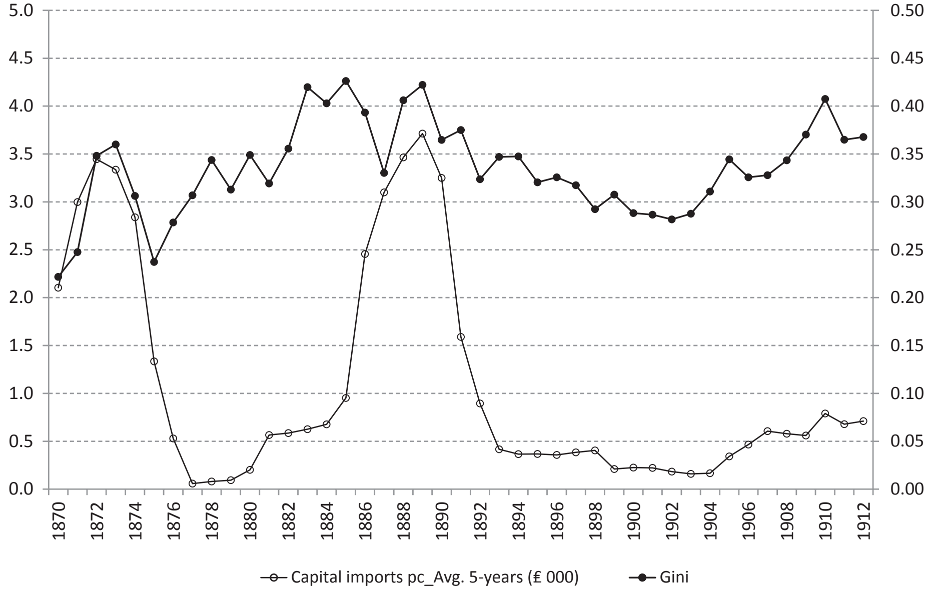

Figure 9. Capital inflows and inequality capital imports from UK (5-year average, per capita, £) and the Gini Index
Source: Stone (1999) and own estimations.

The evolution of the variables deserves an additional comment. During the period, an international shock happened whose economic consequences—and specially in the River Plate–strengthen our argument. The Baring crisis meant a severe recession and, probably, the most intense sovereign debt crisis of the 19th century. This world financial crisis associated to the collapse of the British bank triggered a mayor international crisis. It meant an intense slowdown in world trade and prices, a sudden stop in international capital flows since 1891, and an interruption in migration movements.[9] In other words, the crisis implied a set of anti-globalization forces—or, at least, forces that moderated and rebalanced the structural forces of the globalization—with favorable consequences on inequality (income distribution improved in the 1890s).

Therefore, both in statistical and historical terms, our results offer reliable insights. Our outcomes fulfil the statistical requirements of the estimation and are compatible with the historical facts in the theoretical framework of the H-O model.

5.1. *Robustness Check: Inequality Possibilities Frontier*

Milanovic *et al.* (2011) assume a pre-industrial society which has to distribute income in such a way as to guarantee minimum subsistence level for its poorer classes. The rest of the total income is the surplus that is shared among the richest classes. When average incomes are very low, and close to the minimum subsistence level, the surplus is small. Under these conditions, the members of the upper class will be few, and the level of inequality will be quite small. But as average income increases with economic progress, this constraint is lifted, the surplus can increase, and the maximum possible inequality compatible with that new, higher, average income is greater. In other words, the maximum attainable inequality is an increasing function of mean overall income. Whether the elite fully exploit that maximum, and whether some trickle-down allows the minimum subsistence level to rise, is the matter of economic historians. If we chart the locus of such maximum possible Ginis on the vertical axis against mean income levels on the horizontal axis, we obtain the inequality possibility frontier (IPF) (Figure 10).

How similar are country inequality measures to the maximum feasible Gini indexes at their estimated income levels? The ratio between the actual and the maximum possible inequality is called the inequality extraction ratio, which indicates how much of the maximum inequality was actually extracted: the higher the inequality extraction ratio, the more (relatively) unequal the society.

Are our estimates compatible with the Milanovic *et al.* (2011)'s approach? We considered two aspects of their analysis. First, we plotted our data in Figure 10 to compare them with the rest of the data available. Clearly, in the pre-industrial phase, Uruguay evolved by a trajectory of low inequality compared to its peer countries.

[9]In the 1890s, the inflow of migrants was aggravated by the fact that Spain was the main source (both, for Argentina and Uruguay) and this country suffered a severe depreciation of its currency. For the potential emigrants in Spain, depreciation represented an obstacle. Since the peseta fell in value on average around 30 percent from 1892 to 1905, it can be assumed that emigration costs were 30 percent higher (Sánchez-Alonso, 2000). Thanks to an anonymous referee for shedding light on this point.

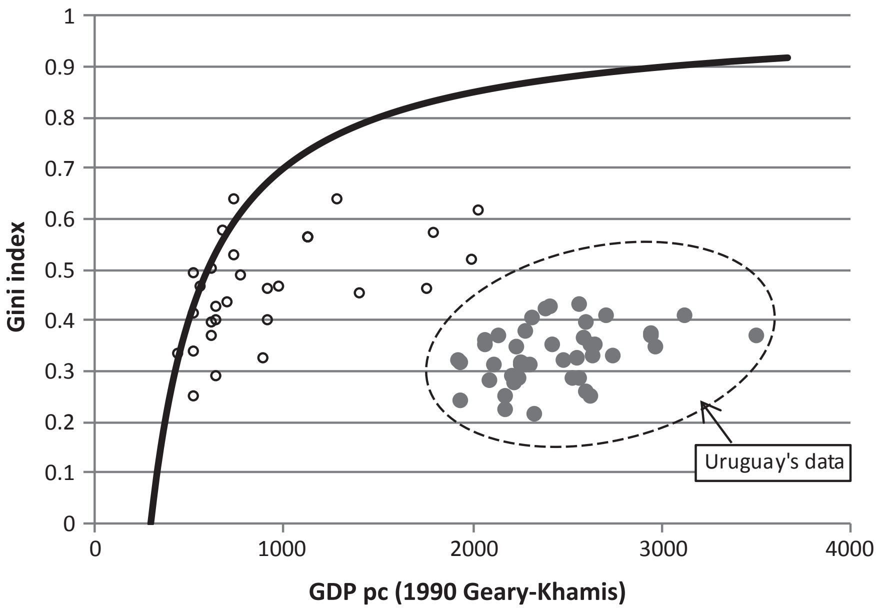

Figure 10. Inequality Possibilities Frontier
Source: Milanovic *et al.* (2011) and own estimates for Uruguay.

This is consistent with the historical evidence and the attractiveness showed by the River Plate as a true "land of opportunities" during the Frist Globalization.

Second, we calculated the extraction ratio and showed the evolution of this indicator in Figure 11. In terms of inequality, this indicator does not add new information and would only confirm that the society was never close to the maximum extraction (the index never exceeded the 50 percent).

However, we face an additional restriction. Is a subsistence level of $300 (expressed in 1990 Geary-Khamis) reasonable? A previous study (Bértola, *et al.*, 2010) dealt with a similar question regarding the Brazilian case. The answer is that people's welfare was not determined for the income expressed in terms of 1990 purchasing power but that corresponding to the current income of the corresponding period. With this objective, we considered a subsistence income equivalent to half of the actual wage. Our wage data represent an average of agricultural wages and we assumed a reduction of 50 per cent to represent that level; this is an absolutely arbitrary decision with the only purpose of obtaining certain order of magnitude. We expressed all wage data in 1912 prices and selected the lower record of the complete series as reference. It corresponded to 1889 and achieved a minimum of $226 (in current prices).

This exercise allows constructing a more "realistic" IPF curve (Figure 12) and the Gini data are significantly closer to the curve than in the more general case (Figure 10).

With this exercise, we show that: (1) our estimates render inequality indicators consistent with the theoretical approach of Milanovic *et al.* (2011) and also coherent with the historical evidence that shows that Uruguay was a country with

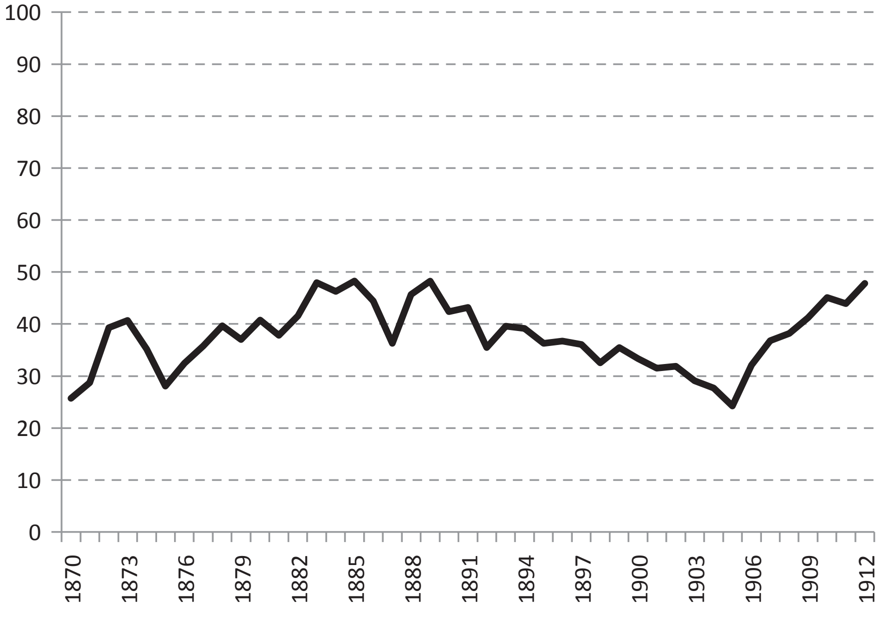

Figure 11. Uruguay: Inequality Extraction Ratios
Source: own estimates.

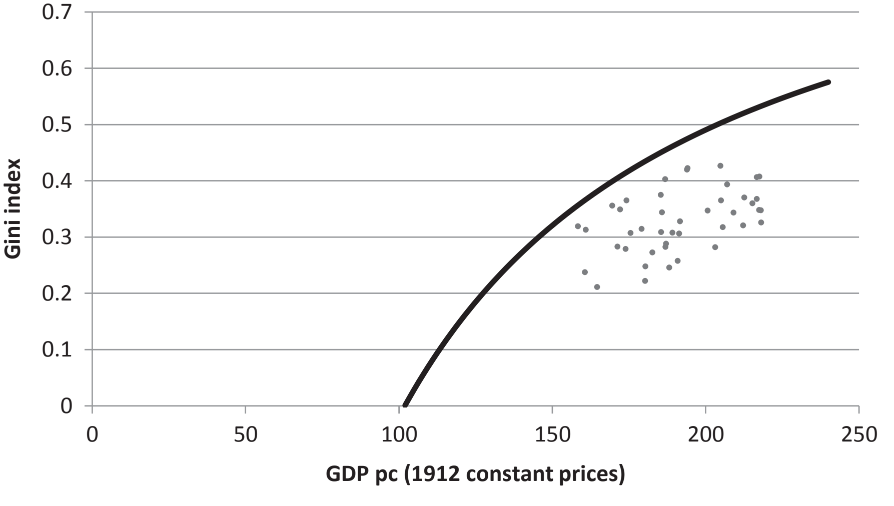

Figure 12. Inequality Possibilities Frontier for Uruguay
Source: own estimates.

relatively lower levels of inequality; (2) From a methodological point of view, this exercise represents additional evidence in favor of questioning the use of "modern" PPPs to evaluate the economic performance of a very remote past (see, for instance, Prados de la Escosura, 2000).

6. Final Remarks

The aim of this paper is to propose different alternatives of inequality estimation for economies with a large agricultural sector where land is a decisive factor in income generation and the available information about personal earnings is not enough. We have data about the incomes of the inferior and superior tail of the income distribution, the GDP, the levels of land rents and wages and the corresponding earners (that we identify with the economically active population).

We propose different exercises considering Pareto and Log-normal income distribution and assuming that the total earners and, alternatively, the land rents, the wages and the GDP are true. According to these statistical considerations, we chose the most feasible estimation which corresponds to a Pareto distribution.

We checked the historical coherence of our estimated series considering the evolution of the main explicative variables of inequality during the First Globalization (according to H-O theory) and Milanovic's framework corresponding to the IPF curve. These exercises confirm the feasibility of our estimates.

Our proposal can contribute to the estimation of inequality indexes in economies similar to Uruguay during the First Globalization where agricultural sector and land resulted decisive in the income generation, and incomes of the inferior and superior tail of the income distribution are available. However, in general terms, the technique applied could be used for all countries that comply with three conditions: (1) data about GDP in current prices is available; (2) data about earners and their respective incomes corresponding to the superior and inferior tails of the distribution (independent on they are landowners, industrialists, businessmen, agrarian or construction workers, etc.) are available; (3) the poorest and the richest people corresponded to the same type of income during the whole period.

References

Álvarez, J., L. Bértola, and G. Porcile, "Introducción". In Álvarez, J., L. Bértola, and G. Porcile (Comp.) *Primos Ricos y Empobrecidos. Crecimiento, distribución del ingreso e instituciones en Australia-Nueva Zelanda vs Argentina-Uruguay*, Ed. Fin de Siglo, Montevideo, Uruguay, 7–53, 2007.

Álvarez-Nogal, C., and L. Prados de la Escosura, "The rise and fall of Spain (1270–1850)," *Economic History Review*, 66, 1–37, 2013.

Barro, R. J., "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth*, 5, 5–32, 2000.

Baten, J., "Economic Development and the Distribution of Nutritional Resources in Bavaria, 1797–1839," *Journal of Income Distribution*, 9, 89–106, 2000.

Bertino, M., and H. Tajam, *El PIB de Uruguay 1900–1955*. Instituto de Economía,Facultad de Ciencias Económicas y de Administración, Montevideo, 1999.

Bértola, L., "A 50 años de la Curva de Kuznets: Crecimiento y distribución del ingreso en Uruguay y otras economías de nuevo asentamiento desde 1870," *Investigaciones en Historia Económica*, 3, 135–76, 2005.

______, *El PIB Uruguayo 1870–1936 Y Otras Estimaciones*, Programa de Historia Económica, Facultad de Ciencias Sociales, Montevideo (with the collaboration of Calicchio, L., Camou, M. and Rivero, L.), 1998.

Bértola, L., L. Calicchio, M. Camou, and G. Porcile, "Southern Cone Real Wages Compared: a Purchasing Power Parity Approach to Convergence and Divergence Trends, 1870-1996". Documento de Trabajo 44, *Unidad Multidisciplinaria, Facultad de Ciencias Sociales, Universidad de la República*, Uruguay, 1999.

Bértola, L., C. Castelnovo, J. Rodríguez Weber, and H. Willebald. “Between the Colonial Heritage and the First Globalization Boom: On Income Inequality in the Southern Cone,” *Revista de Historia Económica-Journal of Iberian and Latin American Economic History*, 28, 307–41, 2010.

Bonino, N., C. Román, and H. Willebald. “PIB y estructura productiva en Uruguay (1870–2011): Revisión de series históricas y discusión metodológica”. Serie Documentos de Trabajo DT 05/12, *Instituto de Economía, Facultad de Ciencias Económicas y de Administración, Universidad de la República*, Uruguay, 2012.

Bourguignon, F., and C. Morrisson, “Inequality among World Citizens: 1890–1992,” *American Economic Review*, 92, 727–44, 2002.

Clemens, M., and J. Williamson, “Wealth Bias in the First Global Capital Market Boom, 1870–1913,” *The Economic Journal*, 114, 304–37, 2004.

Deininger, K., and L. Squire, “A New Dataset Measuring Income Inequality,” *The World Bank Economic Review*, 10, 565–91, 1996.

François, J. F., and H. Rojas-Romagosa, “The Construction and Interpretation of Combined Cross Section and Time-Series Inequality Datasets,” *World Bank Policy Research Working Paper* 3748, 2005.

Lindert, P., and J. Williamson, “Revising England’s Social Tables, 1688–1812,” *Explorations in Economic History*, 19, 385–408, 1982.

Lloyd, C., and J. Metzer, “Settler Colonization and Societies in History: Patterns and Concepts,” in C. Lloyd, J. Metzer, and R. Sutch (eds), *Settler Economies in World History*, Ch. 1, 1–34, 2013.

Lopez, J. H., and L. Servén, “A Normal Relationship?: Poverty, Growth, and Inequality,” *World Bank Policy Research Working Paper Series* 3814, 2006.

Martínez-Galarraga, J., A. Rodríguez Miranda, and H. Willebald, “Patterns of Regional Income Distribution in Uruguay in the Long Run (1870–2008). A Convergence Story”, Ch.12. in D. Tirado, M. Badía-Miró, and H. Willebald (eds.), *Time and Space – Latin American Regional Development in Historical Perspective*, Palgrave Studies in Economic History, Palgrave-Macmillan, London (in press), 2019.

Milanovic, B., *Worlds Apart. Measuring International and Global Inequality*, Princeton University Press, Princeton, NJ, 2007.

______, *Global Inequality. A New Approach for the Age Globalization*, Harvard University Press, Cambridge, MA, 2016.

Milanovic, B., P. Lindert, and J. Williamson, “Pre-industrial Inequality,” *Economic Journal*, March, 255–72, 2011.

______, “Global Income Inequality by the Numbers: In History and Now. An Overview,” *Policy Research Working Paper Series* 6259, The World Bank, 2012.

Millot, J., and M. Bertino, *Historia Económica del Uruguay. Tomo II 1860–1910*, Instituto de Economía-Facultad de Ciencias Económica y de Administración- Universidad de la República, Fundación de Cultura Universitaria, Montevideo, 1996.

Mitchell, B. R., *International Historical Statistics: The Americas 1750–1988*, Stockton Press, New York, USA, 1993.

Moothathu, T. S. K. “Sampling distributions of Lorenz curve and Gini index of the Pareto distribution,” *Sankhyā: The Indian Journal of Statistics, Series B*, 47, 247–58, 1985.

O’Rourke, K. “Globalization and Inequality: Historical Trends,” *NBER Working Papers* 8339, National Bureau of Economic Research, 2001.

Persson, T., and G. Tabellini, “Is Inequality Harmful for Growth?,” *American Economic Review*, 84, 600–21, 1994.

Prados de la Escosura, L., “International Comparisons of Real Product, 1820–1990: An Alternative Data Set,” *Explorations in Economic History*, 37, 1–41, 2000.

______, “Growth Inequality, and Poverty in Latin America: Historical Evidence, Controlled Conjectures,” *Economic History and Institutions Series Working Paper* 05–41(04), Dpto. de Historia Económica e Instituciones, Universidad Carlos III, Madrid, 2005.

______, “Inequality and Poverty in Latin America: A Long-Run Exploration,” in T. J. Hatton, K. H. O’Rourke, A. M. Taylor (eds), *The Comparative Economic History*, MIT Press, Cambridge, MA, 291–315, 2007.

______, “Inequality, Poverty and the Kuznets Curve in Spain, 1850–2000,” *European Review of Economic History*, 12, 287–324, 2008.

Sánchez-Alonso, B., “European Emigration in the Late Nineteenth Century: The Paradoxical Case of Spain,” *Economic History Review, Economic History Society*, 53, 309–30, 2000.

Soltow, L., “The Measures of Inequality,” in L. Soltow, and J. L. van Zanden (eds), *Income & Wealth Inequality in the Netherlands 16th–20th Century*. Het Spinhuis, Amsterdam, 7–22, 1998.

Steckel, R., "Stature and the Standard of Living," *Journal of Economic Literature*, 33, 1903–40, 1995.
Stone, I., *The Global Export of Capital from Great Britain, 1865–1914*. St-Martin's Press, New York, 1999.
Vazquez-Leal, H., R. Castaneda-Sheissa, U. Filobello-Nino, A. Sarmiento-Reyes, and J. Sanchez Orea, "High accurate simple approximation of normal distribution integral," *Mathematical Problems in Engineering*, 2012, Article ID 124029, 22 pages, doi: https://doi.org/10.1155/2012/124029. 2012.
Willebald, H., *Natural resources, settler economies and economic development during the first globalization: land frontier expansion and institutional arrangements*. Universidad Carlos III de Madrid, Departamento de Historia Económica e Instituciones, Getafe, http://biblioteca.uc3m.es/uhtbin/cgisirsi/g5tdKk6Xgq/CCSSJJ/83570128/123. 2011.
______, "Distributive patterns in settler economies: agrarian income inequality during the First Globalization (1870–1913)," *Historia Agraria. Revista de Agricultura e Historia Rural*, 66, 75–104, 2015.
Williamson, J., "Globalization, Factor Prices and Living Standards in Asia before 1940," in A. J. H. Latham, and H. Kawakatsu (eds), *Asia Pacific Dynamism 1550–2000*, Routledge, London, 13–45, 2000a.
______, "Real Wages and Factor Prices Around the Mediterranean 1500–1940," in S. Pamuk, and J. G. Williamson (eds), *The Mediterranean Response to Globalization Before 1950*, Routledge, London, 45–75, 2000b.
______, "Land, Labour and Globalization in the Third World, 1870–1940," *Journal of Economic History*, 62, 55–85, 2002.
______, "Real Wages, Inequality, and Globalization in Latin America Before 1940," *Revista de Historia Economica*, 17, 101–42, 1999.
van Zanden, J. L., J. Baten, P. Foldvari, and B. Leeuwen, "The changing shape of global inequality 1820–2000; exploring a new dataset," *Review of Income and Wealth*, 60, 279–97, 2014.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site.

Appendix 1. Lognormal vs Pareto Estimates of GDP
Appendix 2. Our Estimates and Comparison with Williamson Index
Appendix 3. Statistical Data