



Inequality in late colonial Indonesia: new evidence on regional differences

Pim de Zwart¹

Received: 4 August 2020 / Accepted: 25 November 2020
© The Author(s) 2021

Abstract

This paper adds to a growing literature that charts and explains inequality levels in pre-industrial societies. On the basis of a wide variety of primary documents, the degree of inequality is estimated for 32 different *residencies*, the largest administrative units and comparable to present-day provinces, of late colonial Indonesia. Four different measures of inequality (the Gini, Theil, Inequality Extraction Rate and Top Income Rate) are employed that show consistent results. Variation in inequality levels across late colonial Indonesia is very large, and some residencies have much higher levels of inequality (with, for example, Ginis above 60) than others (with Ginis below 30). This suggests that even within a single colony, levels of inequality may vary substantially and this puts some doubts on the representativeness of using a single number to capture the level of inequality in a large economy. In order to explain the variation across residencies and over time, this paper investigates the role of exports and plantations, so frequently mentioned in the literature. It is shown that both explain a part of the variation in levels of inequality across colonial Indonesia, but that only the rise of plantations can explain changes in inequality levels over time. This points to the importance of the institutional context in which global export trade takes place for the rise of inequality.

Keywords Southeast Asia · Inequality · Colonialism · Globalization

JEL Classification F6 · I3 · N3 · N5 · N7

1 Introduction

Over the past years, the issue of within-country economic inequality has surged as a topic of interest among economists and economic historians (e.g. Piketty 2014; Milanovic 2018; Lindert and Williamson 2017). Charting the global patterns of

✉ Pim de Zwart
pim.dezwart@wur.nl

¹ Wageningen University, Wageningen, Gelderland, The Netherlands

within-country inequality and investigating the potential drivers of these patterns over time are a crucial task for economic historians. Despite the growing interest, the evidence documenting inequality in developing countries pre-dating the 1960s remains scarce (only 20 data points for Ginis in the pre-1960s Global South were recently enumerated by Lindert and Williamson 2017: 292–293). This study contributes by gauging the level of income inequality in Indonesia, then known as the Dutch East Indies, in the early twentieth century. Indonesia was, after India, the largest and most populous European colony and one of the main producers of the tropical commodities (such as rubber and sugar) that fuelled the global economy in this period. The case of Indonesia thus allows to analyse how inequality is affected by global trade within a colonial institutional context. This is particularly important as both the level of inequality and its effects for economic development are influenced by the wider institutional setting (Bennett and Nikolaev 2017; Holcombe and Boudreaux 2016). Some recent research suggests that economic development is negatively affected by high levels of inequality when economic freedom is lacking (Apergis et al. 2013), such as under colonialism.

This is not the first study of inequality in colonial Indonesia. Geertz (1963) famously suggested that the Javanese in the nineteenth century suffered from fast population growth that led to decreasing inequality and “shared poverty”. Investigations by Booth (1988, 1998) have highlighted a high degree of inequality, especially between different ethnic groups (Indonesians, Chinese and Europeans) in the late colonial period. van Zanden and Marks (2012) note a substantial increase in inequality between 1880 and 1925 as the Gini rose from 0.39 to 0.48. A study by Leigh and van der Eng (2010) examined the share of total income that went into the pockets of the top 1% of Indonesian society. They find that the top 1 percent obtained 12 percent of the total income in 1921, which increased to over 22 percent in 1933–1934 and then declined slightly to just under 20 percent in 1938–1939. In a recent study, van Leeuwen and Földvári (2017) found high and increasing levels of inequality from the 1920s to the late 1930s with Ginis of around 0.50. Additionally, they find that the share of Indonesians living below the poverty line increased substantially over the colonial era.

In terms of explaining the trends in inequality, many of these authors emphasize the role played by global export trade. Leigh and van der Eng (2010) suggest that increasing top income shares in the 1920s and early 1930s was the result of drops in global commodity prices that hit smaller farmers in Indonesia more than the top income earners, who were largely shielded from such developments. van Leeuwen and Földvári (2017: 387) note that rising inequality was possibly caused by increased land rents due to the growth of cash crop production for the global market. They build on the argument made by Williamson (2011, 2015) for colonial Southeast Asia more broadly who argued that in land abundant countries increased global trade tended to benefit landholders rather than labourers. Most recently, Booth (2019: 104) confirms this line of thinking and notes that, for example, in Java the ratio of rent to wage payments clearly rose throughout the 1920s and 1930s.

None of these studies, however, have investigated how inequality varied between different parts of the archipelago, but rather analysed trends for Indonesia as a whole. Yet because Indonesia is such a large (not much smaller than Europe in its

entirety) and diverse country, differences in both levels and trends of inequality at the sub-national level may be substantial (see also recent work by Bosma 2019). In addition, studies have pointed out the divergent impact of Dutch colonialism in Java compared with the so-called Outer Islands of Indonesia (i.e. Sumatra, Borneo, Bali, Lombok, Celebes, New Guinea, Timor, Ternate, and the Moluccas) (Foldvari et al. 2013). Others noted the differences in colonial rule and patterns of development across Java (e.g. Elson 1994), and between the various regions in the Outer Islands (Touwen 2001).

In this article, I put forth new estimates of inequality for colonial Indonesia in the first half of the twentieth century. In doing so, the focus, in contrast to the studies mentioned above, is not to sketch the long-run trends in inequality, but to find out to what extent the degree of inequality differed between the various regions of Indonesia. Regional disaggregation of inequality estimates allows to study which factors have influenced the income distribution (Lindert and Williamson 2016; Geloso 2018). Many of the underlying conditions that have been put forward to explain inequality in the past, such as interaction with the global market (Williamson 2011) and the institutional framework in terms of the plantation system (Sokoloff and Engerman 2000), as well as urbanization and population densities (Milanovic 2018), varied greatly across the archipelago. Therefore, in this paper, the unit of analysis will be the *residency*, which was the administrative level below that of the state and more or less comparable with present-day provinces. A second aim of this paper is to consequently explain some of the variation in inequality levels across the Indonesian archipelago by looking at the role of export trade and the plantation system.

This paper proceeds as follows: in Sects. 2 and 3, I will make use of colonial income tax data from the 1920s, evidence on land distribution, agricultural production and wages in the different residencies, in combination with evidence from a number of investigations on the welfare of the indigenous population that were issued at the time, in order to estimate the degree of inequality in the different residencies of Indonesia. In Sect. 4, possible factors influencing the differences in inequality will be examined and it is shown that plantation agriculture is one of the main drivers of high and increasing inequality. Section 5 will offer the main conclusions from this investigation and provides an agenda for further research.

2 Computing Ginis

The main source in this paper concerns the tabulations of the colonial income tax as they appeared in the Colonial Reports (*Koloniale Verslagen*) from 1922 to 1930 (referring to the taxation in the years 1920 to 1928). These sources have also been used by Leigh and Van der Eng (2010) to analyse the developments in the income share of the top 1% for Indonesia as a whole. A general income tax was implemented in 1920 to which all income earners and companies in Indonesia were liable. Income earners from the same households were taxed together. Assessments for the income tax were made by village (*desa*) heads rather than by the colonial officials. Reys (1925: 75) notes that as the village heads were chosen by the village population and tax assessments were being prepared in collaboration with the taxpayers, it is likely

that they underreport the actual incomes to be assessed. In particular among the lower income brackets underreporting seems to have been a serious problem (Leigh and Van der Eng 2010). This means that the figures shown here possibly understate the true level of inequality.

Enforcement and monitoring of the tax by Dutch colonial authorities were limited (Reys 1925: 76–77). A recent study on variation in enforcement of the twentieth-century income taxes in the USA has shown that this impacts inequality estimates (Geloso and Magness 2020). Differences in enforcement across the Dutch East Indies may thus drive a bias in the inequality estimates for the separate residencies. It is possible that limited capacity for enforcement led the colonial state to focus its enforcement efforts on wealthier regions, or regions with more Europeans. In those regions, greater enforcement could have led to more accurate reporting at the highest income brackets, thus increasing inequality estimates.¹ Unfortunately, there are no detailed, regionally disaggregated, reports on tax enforcement, so it will be hard to come to any robust conclusions regarding this issue, yet it is important to keep in mind when assessing the figures later.

The tax assessments showed figures separately for different ethnic groups: Europeans, “Foreign Asiatics” and Indonesians. Leigh and Van der Eng (2010: 177) note that these groups are misleading, as people were not necessarily classified according to ethnic background: Chinese, for example, could be noted down as both “Foreign Asiatics” or as “Europeans”, while Indo-Europeans as “Europeans” or as “indigenous”. These categories have therefore been omitted from this analysis. Instead, all colonial legal categories have been collapsed into one figure for each residency.

The overall coverage of this tax was exceptionally high for a developing country in this period. In 1920 already 2.6 million people were paying the income tax (representing 22 per cent of households), which increased to over 4 million in 1930 (30 percent of households) (Leigh and Van der Eng 2010: 177). The coverage differed substantially between the different residencies, however. For some residencies, coverage of the tax was about 100 per cent (e.g. in West Sumatra, Jambi and Palembang), whereas for others this could be about 10 percent (in the residencies of Kedu, Rembang and Madiun in Java) and others somewhere in between those figures.² The reason for this difference is that incomes below fl. (guilders) 120 annually were exempted, as were those people liable to the land tax. It seems likely that the greater extent of poverty among the population in Java implies that fewer people had to pay the income tax there. Land taxes were only levied in Java, Bali, Lombok and

¹ Geloso and Magness (2020) find that more aggressively enforced income taxes may reduce inequality estimates, as enforcement efforts are then more spread also through the lower brackets of the income distribution. On the basis of Reys (1925) and Leigh and Van der Eng (2010) it seems more likely that in colonial Indonesia, more enforcement in certain regions will have been focussed towards the top incomes. Without good comparative data or more detailed information, it is hard to come any final conclusions on this.

² Since members of households were tax together, the records in the income tax relate to households. In order to estimate the coverage per residency related to the total population it was assumed that households everywhere consisted on average of five persons. This is a rough estimate, but gives some indication of coverage. See also Leigh and Van der Eng (2010).

Southern Celebes (Booth 1980: 102). An additional reason for low coverage may be the problem of underreporting.

The fact that low coverage is generally more of a problem for Java than for the other islands (e.g. Sumatra is rather well covered by the income tax data) is perhaps somewhat surprising as due to the longer history of colonial rule and generally more extensive colonial bureaucracy in Java one would expect data issues to be less problematic there. For this research, I only employ figures for 1920, 1924 and 1928 (from the KVs of 1922, 1926 and 1930), as after that data on the income tax that appeared in the Indian Reports (*Indische Verslagen*, continuation of the KV) reported only tax information for Java and the Outer Islands as a whole and were therefore unsuitable for computing inequality in each of the separate residencies.

With low coverage for some of the residencies in the sample it is necessary to correct for this underreporting and add estimates of incomes for those not included in the income tax records: that is (1) households paying the land tax, and (2) households with incomes below the fl. 120 threshold. Because the income tax was levied at the household level, in this paper inequality is calculated between households, rather than the total population. To do so, the model to compute household income derived from land is employed:

$$Y_{h,i} = H_{h,i}R_iP_i + W_i(L_{h,i} - 210H_{h,i}) \quad (1)$$

where Y is total income from land per household (h) in each residency (i), H is hectares of land held by the household, R the average production of paddy per residency, P the price of paddy per residency, L is total amount of labour days per household in each residency and W is the average unskilled wage in a residency.

Information about the amount of hectares land owned per household is taken from a large-scale investigation into the welfare of the Javanese population conducted at the beginning of the twentieth century, known as the *Declining Welfare Study* (Hasselman 1914: Appendix R).³ The reports from this investigation show the number of households holding land of varying sizes in Java in 1903. Unfortunately, there are no data on the distribution of land in the 1920s, so it has to be assumed that land inequality does not change much between those years. It is, however, likely that land inequality increased in this period, as has been discussed by various scholars (Wertheim 1956; Boeke 1955; Scheltema 1927), and was also noted in the report on tax pressure on the Javanese in the 1920s (Meijer Ranneft and Huender 1926). This means that the estimates in this paper possibly understate the true level of inequality in Java.⁴ In Appendix 1, it is analysed how data on land distribution in the 1920s might have altered computed Ginis. It is shown that even taking different assumptions regarding changes in land distribution on the basis of some of the information we do have, this has no effects on the main results of this paper.

³ *Onderzoek Mindere Welvaart*, henceforth OMW.

⁴ Considering the conclusion of relatively high levels of inequality in Java, as well as the conclusions that the share of planted estate area (which is relatively high across Java) are highly correlated with the degree of inequality, these assumptions are biased against my conclusions. More complete data would increase inequality in Java.

In order to translate this information on land holding into monetary incomes, we add to this information on average production and estimate the costs of labour inputs. To do so, it was assumed that only rice was being produced on these lands. While we know that other crops were also being cultivated, most of the land was devoted to rice and it is a safe assumption to get an indication of the income that could be derived from such lands. Average production per hectare is obtained from Boomgaard and Van Zanden (1990) who provide figures on the total production of paddy (unhusked rice) in the Java residencies, as well as the area planted with paddy, thus allowing the calculation of the average amount of paddy produced per hectare across the Java residencies. Market prices for paddy in all the different residencies were reported for 1928 (van Zanden and Marks 2020). As market prices of paddy were only available for Jakarta in 1920 and 1924, these were used to estimate rice prices for the other residencies assuming similar price differentials as in 1928 (see Appendix 2 for more details).

In order to produce these quantities of rice, it is assumed that this required substantial inputs of labour, but no capital. For Java in the 1920s it was estimated that 1 hectare of land required a labour input of 210 working days (van der Eng 2004). This concerns inputs on irrigated rice fields (*sawah*) that were common in Java, yet labour inputs may have been somewhat lower on dry fields (*ladang*) that were more common in most other islands. Lacking estimates of such inputs across the archipelago, it was decided to use the Java estimate for the East Indies as a whole. The Dutch East Indies Census of 1930 (EZ [Dept. van Economische Zaken] 1936: 126–127) provides information about on labour market participation rate, which ranges from 0.29 in West Java to 0.36 in East Java (as a share of the total population).⁵ If we would assume a total of 250 days per worker (Allen et al. 2011), this results in a total of 365–456 days of labour per household depending on the labour market participation rate. Of course, this excludes labour performed that was out of view from the census takers, so that participation, for example, by women and children, may be higher than suggested in the census data. Unfortunately we have no estimates about this across the various regions, but we should keep in mind that we underestimate contributions by women and children (e.g. van Nederveen Meerkerk 2019: 184). For now, in order to estimate household incomes, it is calculated on the basis of these figures that households can work about 2 hectares of land themselves without having to hire additional labour. Households with smaller plots of land than 2 hectares can use their remaining days to work for wages. This wage income was added to the income from the land. Households with greater plots, on the other hand, would have to hire wage labourers to work their fields. Wages for the period 1920–1924 were quoted by the report of Malines van Ginkel (1926). No wages for 1928 could be found there and therefore wages for 1924 were extrapolated using the data on labourers in the sugar industry from Dros (1992) (see Appendix 2 for details). Combining this information with the figures on the sizes of plots (introduced above) allows for the calculation of household incomes derived from land

⁵ This participation rate is a percentage of the total population and not of the people of working age.

(Appendix 5 gives an overview of the data and sources used to estimate income from land in Java in 1920, 1924 and 1928).

Finally, in order to compute a Gini (the most common measure of inequality, which ranges from 0, implying perfect equality, to 100, perfect inequality), other households not covered by the income tax are assumed to be those earning below fl. 120. However, since this is also roughly the price for subsistence, it was assumed that incomes could not have been much below that figure. Therefore, those households have been assigned fl. 115 as annual income. In Appendix 3 estimates are shown using other assumptions regarding the subsistence income, demonstrating that it hardly has an effect estimated Ginis and does not affect the results of the regressions analyses shown in Sect. 4. The number of households that have been assigned this income level differs per region, of course, depending on how many households were already included in the income tax and the number of households that already showed up in the records as holding land. Since we may also assume some underreporting in the various income brackets, while the number of landholding households may also have increased somewhat between 1903 and 1920, we include so many households earning fl. 115 in each region so that the total number of households covered by the “corrected” Gini is at least 90%. Using this yardstick, for Java, some 40 percent of all households fall into this “subsistence” category, while for the Outer Islands this figure is some 25 percent.

The computations on the spread of incomes across these different Indonesian residencies are based on a number of assumptions. For this reason, I will introduce another measure of inequality (the Top Income Rate, or TIR) that does not require these same assumptions in the next section. For now, it is important to see how the “corrected” figures, with adding information about incomes from land and assigning subsistence incomes of fl. 115 to the remaining population, affects the results. Figure 1a–c shows the correlations between the Gini among income taxpayers and the corrected Ginis using data for 1920, 1924 and 1928, respectively. Since more alterations were made to the Ginis for Java (leading to a greater reduction in inequality) than for the Outer Islands, these are shown as separate series. Figure 1 makes clear that the cross-sectional correlation between the original and corrected Gini is quite strong. As the corrected Gini, in theory, measures the level of inequality between more or less the entire population of a residency (90% or above), those are used in the remainder of this paper.

Both income taxes and land taxes levied may have affected the distribution of income. The income tax was a progressive tax, and therefore, taxation may have impacted the distribution of income in these regions. From the *Staatsblad* (1920: Art. 678), which contains the income tax laws, one can discern the different income tax brackets, the percentage of tax these had to pay over their income, as well as the lump sum related to the different tax brackets. In this law, the lowest income tax rate was 1 percent on incomes below fl. 120 and a lump sum of fl. 1.20 (which only had to be paid if total income was above that), increasing to 2 percent on incomes between fl. 120 and fl. 1800, 3 percent on incomes above fl. 1800 up to fl. 3600, and so on, up to 11 percent on incomes of fl. 36,000 and higher (as well as a lump sum of fl. 2566.80). Next year, additional income tax rates are introduced for the income brackets above 36,000 guilders per annum (*Staatsblad* 1921: Art. 312). This scale

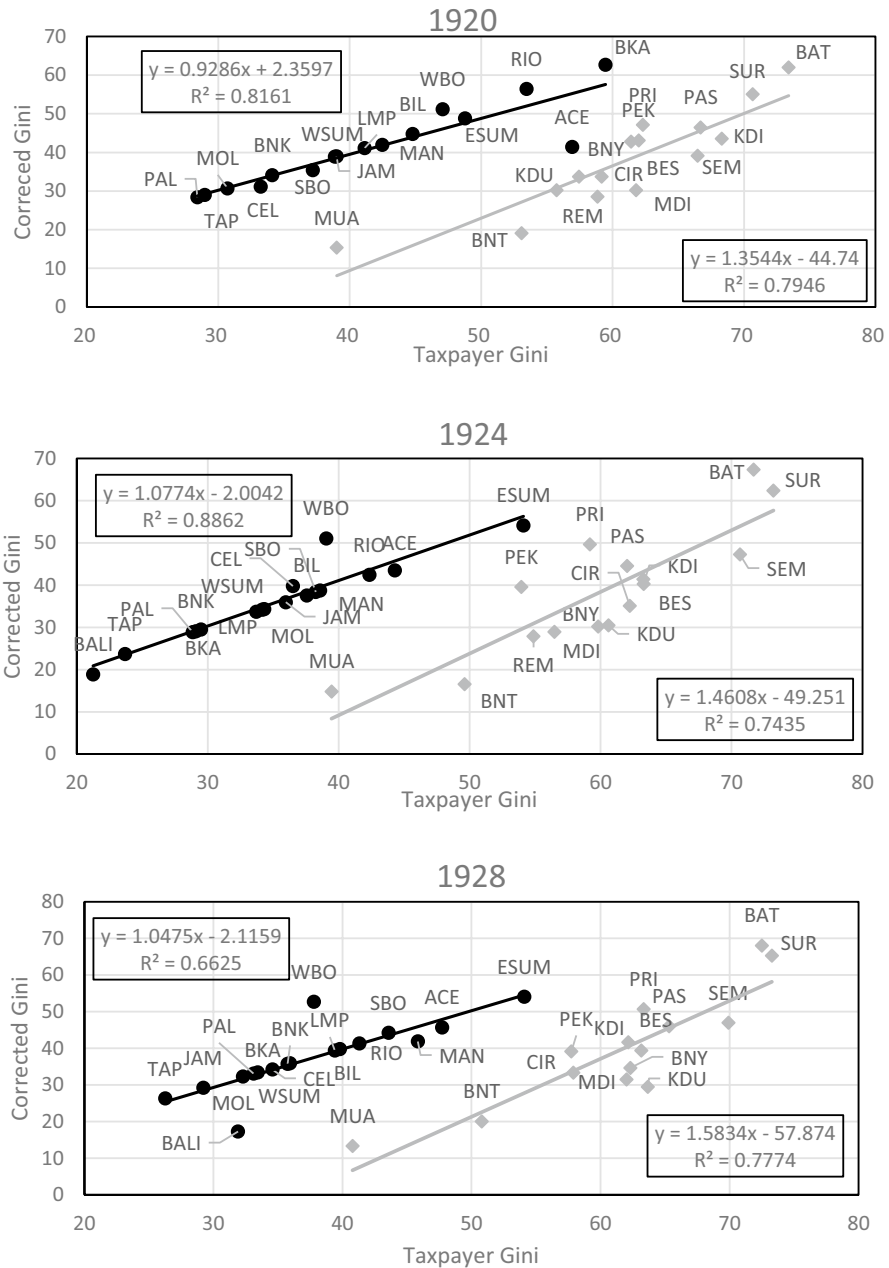


Fig. 1 Correlation between Gini among taxpayers and corrected Gini: **a** 1920, **b** 1924 and **c** 1928. *Sources:* See text. Black dots are residencies in the Outer Islands and grey diamonds are residencies in Java

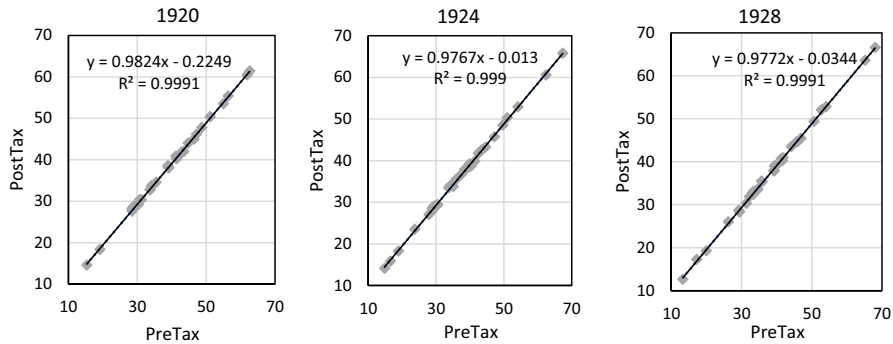


Fig. 2 Correlation between pre-tax and post-tax Ginis, 1920, 1924 and 1928. *Sources:* See text

reaches up to 25 per cent for all incomes earned over fl. 180,000 and a lump sum of fl. 25,966.80. For all different income brackets reported in the income tax records, the amount of taxes due were calculated and deducted from the assessed incomes in order to generate after-tax Ginis.

Land taxes differed according to location. Based on data about location and climate conditions as well as other factors, the colonial tax collectors created three different tax districts based on the “state of economic development”, for districts classified as “low development”, the tax rate was 8–11 percent, for “average development” tax rates of between 12 and 16 percent applied, and for districts with “high development” rates were between 17 and 20 percent—although in practice a maximum of 18 applied. These percentages were then levied on the total paddy produced per *bouw* minus 10 picul of paddy, but with a minimum of the monetary value of 2 picul paddy per *bouw*.⁶ So on a field of 1 *bouw* producing 30 *picul* of paddy in the “low development” district the tax would be the monetary value of 8–11 per cent of 20 *picul* paddy (Tjhan 1933).⁷ Contemporaries have suggested that the tax was in practice regressive as fields that were highly productive were taxed only marginally higher (see Wellenstein 1926; Tjhan 1933). Further problems arose when unproductive lands were situated in districts classified as highly developed. In addition, tax percentages did not increase with the amount of land, so that large landowners were not taxed heavier than those with small amounts of land. In order to obtain information about post-tax Ginis in this paper, we have to make due with information about the land tax at the aggregate level of the residency. Average amounts of land taxes gathered per *picul* paddy in the various residencies were obtained from a report on tax pressure in Java (Meijer Ranneft and Huender 1926: 34), and were deducted from the paddy price to estimate post-tax incomes from land.

The correlation between the pre-tax and the post-tax figures is, as expected, near perfect (Fig. 2). Across the years, it seems that after-tax Gini coefficients are some 2 percent lower, so taxation reduced inequality only very marginally. In the remainder

⁶ One *bouw* is 0.71 hectare.

⁷ One *picul* is 61.76 kg.

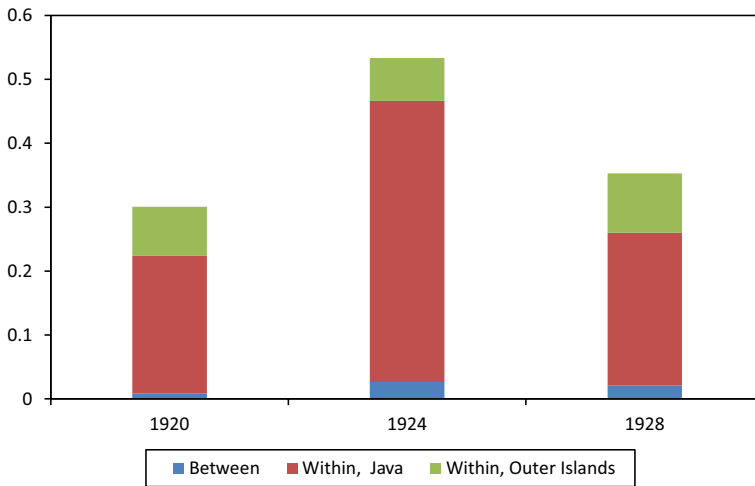


Fig. 3 Theil of inequality across Indonesia, between and within residencies, 1920–1928. *Sources:* See text

of the paper, only the pre-tax figures will be used, as we also want to find out what factors influenced inequality other than the tax system. Considering the correlations shown in Fig. 2, it is clear that there will be no difference in the results of any analysis using pre- or post-tax figures.

3 Additional measures of inequality

Besides the Gini coefficient, inequality can also be measured using the Theil index. The Theil index is more sensitive to income differences at the top end of the distribution, compared with the Gini. The same information discussed above can also be used to compute the Theil ratios in the different residencies using Eq. 2:

$$T = \sum_i w_y^i \log(w_y^i / w_p^i) \quad (2)$$

where T is the Theil index, i is income bracket, w_y is the share in total income and w_p is the share in total households.⁸ The Theil index has the added advantage that it can also be used to decompose inequality across Indonesia. To what extent is the total inequality figure for Indonesia as a whole, used often in previous research, driven by inequality across population groups and to what extent is it driven by differences in distributions between regions? Figure 3 shows the total Theil index for the Dutch East Indies, and what part is driven by within-region inequality and what part by between-region inequality. From the figure it becomes clear that a large part, some 95 percent or more, of total income inequality in the Dutch East Indies in the 1920s

⁸ I follow Frankema and Marks (2010: 82) on computing the Theil index.

Table 1 Prices for a barebones consumption basket in Batavia. *Sources:* Text and Appendix 5

	Unit	Quantity p.p. p.a.	Nutrient per kg		Nutrients per day		Annual costs in <i>fl.</i> (guild- der)		
			Calories	Protein	kcal	Protein	1920	1924	1928
Rice	kg	100	3620	75	992	21	13.73	9.00	8.45
Soya beans	kg	20	4460	360	244	20	5.02	3.39	2.88
Cassava	kg	165	1590	14	719	6	6.14	3.39	2.46
Sweet potatoes	kg	55	860	160	130	24	1.96	1.08	0.97
Fish	kg	6	1301	192	21	3	3.72	3.30	2.91
Salt	kg	3					0.39	0.39	0.39
Lamp oil ^a	litre	1.2					0.72	0.43	0.44
Cotton	m2	3					2.46	2.21	1.06
Petrol	litre	12					2.52	2.52	2.28
Totals					2105	74	36.66	25.72	21.85
HH cost of living (incl. rent)							149.95	105.18	89.36

^aCoconut oil

is the consequence of *within residency* income inequality. Most of the inequality can be found within residencies in Java, but a good 20 or 25 percent of total inequality is generated in the Outer Islands. In terms of trends over the 1920s, the overall Theil index is shown to increase between 1920 and 1924, primarily due to rising inequality in Java, and after that it decreases towards 1928. In order to understand the causes of inequality in the Dutch East Indies, we thus need to understand what drove inequality among the various population groups within the residencies, and why (parts of) Java were so much more unequal than many regions in the Outer Islands. This will be done in the next section, after we have introduced two more measures of inequality.

In order to take account of the fact that in pre-industrial societies with low incomes (close to subsistence) there is a maximum level of inequality, Milanovic et al. (2011) developed the “inequality extraction ratio”. They define an inequality possibility frontier where the maximum Gini=0 in a society with average incomes at subsistence level, but rises to about 60 when average incomes are 2.5 times subsistence level. In order to define what income reflects subsistence level in Indonesia in the 1920s, a consumption basket was created. The basket is inspired by the methodology pioneered by Allen (2015, 2017), who defines a so-called barebones basket that contains a minimum of 2100 calories per day as well as some necessities in terms of lighting, fuel, and clothing. For Indonesia, the basket was created according to these principles, but combined with information on the “cost of some important articles consumed by the native inhabitants of Java” from published sources of the Dutch East Indies agricultural department (LNH [Dept. van Landbouw, Nijverheid en Handel] 1928a). Price data were gathered from separate publications on prices of various commodities (CKS [Centraal Kantoor voor de Statistiek] 1927, 1931).

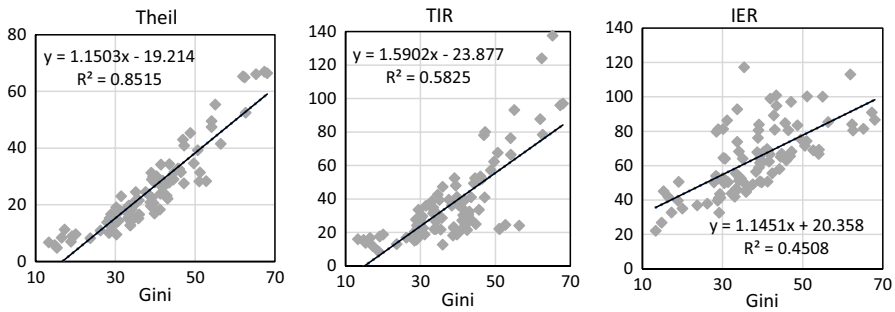


Fig. 4 Correlations between Gini and different measures of inequality: **a** Theil, **b** TIR, **c** IER. Sources: See text and Appendix 5

The basket and the annual costs incurred on its various contents in Batavia (present-day Jakarta) are shown in Table 1. To calculate total household cost of living the cost of the basket was multiplied by 4.09 to account for other family members and house rent. This is a slightly smaller figure than the 4.2 multiplier Allen (2015) uses, because rents were lower in colonial Indonesia.

As became clear from the previous section, the data used to compute inequality for all layers of society are not without problems and assumptions had to be made about the income of those households not included in the income tax records. One way to avoid these assumptions is to look at the incomes of those at the higher ends of the income distribution, as we may believe that the incomes of those people are assessed more accurately than those of the entire population. This is what Piketty (2014) did in his seminal study, and what Leigh and van der Eng (2010) have done for Indonesia as a whole for the period 1920–2004. They have, however, estimated the share of the total income this top 1 percent accrued. While we do have estimates on the total income for Indonesia as a whole for the nineteenth and twentieth centuries (van der Eng 2010), we do not have such information for the separate residencies. While we could rely on total income as estimated in the previous section, this would mean relying on some of the same assumptions again. Instead, we can also estimate the average income of the top 1 percent of the population and relate this to prevailing levels of unskilled rural wages, which can perhaps be seen as representative for the incomes of those at the bottom of the income distribution. We can thus compute the Top Income Ratio (TIR) as an additional measure of inequality.

Figure 4a–c shows the correlations between the Gini, Theil, IER and the TIR for all years (1920, 1924 and 1928) and regions. It becomes clear that there is a strong association between the Gini and other measures of inequality. As expected, the Theil and Gini show the strongest association, as they are based on the same information, but even when including different types of information (e.g. wage and price information) and assumptions to the data, the measures of inequality are still related. We would not have expected the correlation to be perfect, since these measures capture different elements of inequality: with the TIR capturing only the top and the bottom, while the IER takes into account variations in average incomes and prices across these regions. On average, however, regions with a high Gini will also have a

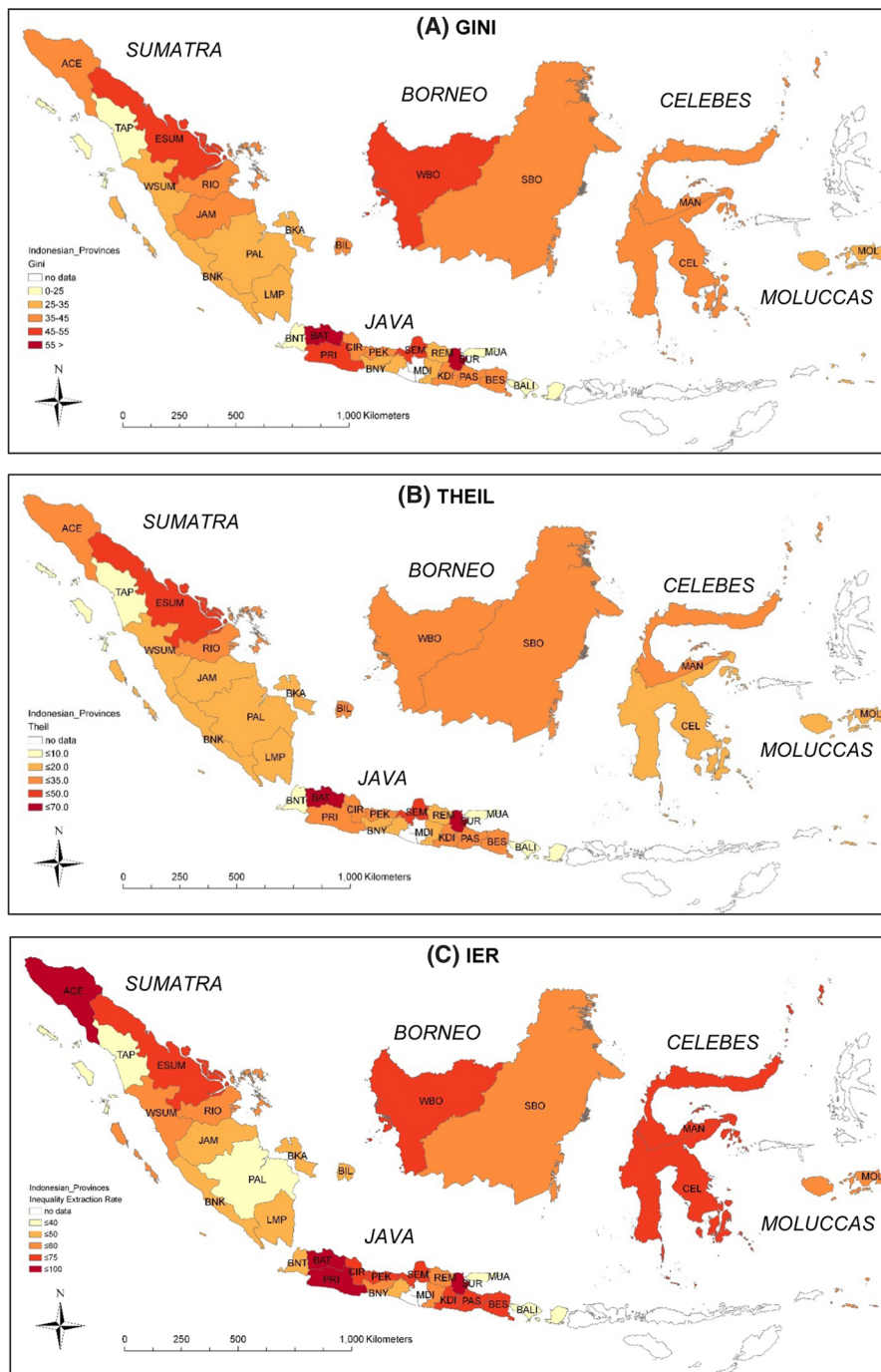


Fig. 5 a Gini, b Theil, c IER, d TIR, in Indonesia in 1924. Sources: See text. Residency borders from Cribb (2000)

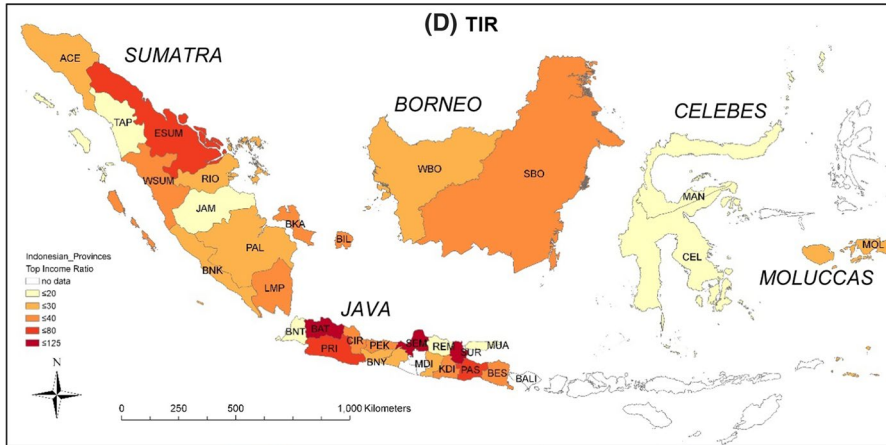


Fig. 5 (continued)

high Theil, TIR and IER. In the analyses later in this paper, these different measures will be used in robustness tests when examining the factors that may have influenced the level of inequality.

In order to get a picture of how inequality differs geographically, Fig. 5a–d shows a map of the Dutch East Indies with the levels of inequality in each residency using data from the 1924 income tax records. Looking at the Gini (Fig. 5a), inequality is highest in Batavia (BAT), Surabaya (SUR), Priangan (PRI), Semarang (SEM), East Sumatra (ESUM) and West Borneo (WBO). More or less a similar picture emerges from Fig. 5b depicting the Theil index in the different residencies. Batavia, Surabaya and Semarang were the regions with the largest European presence. These were also the residencies with large cities where the wealthiest Europeans, Chinese and Indonesians settled. As a result, these may also be the regions where most of the colonial tax officials' enforcement efforts are focussed, which may drive up recorded inequality *vis-a-vis* some of the other regions, due to fuller reporting of incomes at the top income levels. Recent studies have also noted the extractive nature of local governance in Priangan leading to high levels of inequality (Breman 2010; de Zwart 2020). For East Sumatra, the dominance of plantation agriculture and low incomes among unskilled plantation labourers and ensuing inequality have long been noted in the literature (Kian-wie 1969; Pelzer 1978; Breman 1989).

More in the middle range of inequality are the regions of South East Borneo (SBO), Celebes (CEL), and other regions in Java and Sumatra. Lower levels of inequality can be found for regions like West Sumatra, Jambi and Rembang, while the lowest inequality is reported for Banten, Madura, Bali and Tapanuli. Low inequality can be found in regions where export expansion was driven by smallholders or in regions with relatively low average incomes. For residencies of Jambi and West Sumatra, relatively low inequality fits with the literature on the importance of indigenous dynamics in the export economy, rather than large-scale European plantation agriculture (e.g. Touwen 2001; de Zwart 2020). In areas with low incomes and

where the economy was characterized by subsistence agriculture and near-complete absence of an export economy, inequality was also low. There was relatively little export trade from Benkulu, Tapanuli and Bali and what little export production did take place was in the hands of indigenous smallholders (Touwen 2001: 91–93). Low inequality in Banten may be related to the low levels of average incomes there. Poor soils hindered the rise of a thriving export economy, even while located nearby the main trading port of Batavia. It features prominently Multatuli's *Max Havelaar* (1860) as the poor region where a local district head appropriates peasants' buffalo, because there is nothing else to extract. A similar story can be told for Madura, where unproductive lands led to low average incomes and hardly any export activity at all (Bosma 2019: 80–81).

Looking at the IER slightly alters the picture, as various residencies in Java have high extraction rates due to lower average incomes and/or higher prices. Lower average incomes across Java (and thus presumably higher poverty rates) are the result of the much higher population densities there and is consistent with the literature.⁹ Additionally, relatively low average incomes and high price levels in Aceh and across Celebes cause the IER to rise in those regions. On the other hand, Palembang joins the low inequality group as it has relatively high average incomes (much of it earned in rubber and petroleum production), which are apparently relatively widely spread through the population. Finally, the Top Income Rate shows more or less similar patterns. The main difference is that top incomes are also relatively low in Celebes, Tapanuli and Jambi: provinces with little European economic activity (Touwen 2001).

There is no clear geographic clustering of inequality pockets: for example, in Java, regions with very high inequality (such as Batavia and Priangan) were neighbouring regions with very low inequality (Banten). The same can be observed in Sumatra, where egalitarian Tapanuli is located next to highly unequal East Sumatra. The lack of a clear geographical clustering also means that no one clear factor underlying this income distributions emerges easily from this map. High inequality can be observed in both regions with high population densities (Batavia, Surabaya and Semarang) and low(er) population densities (East Sumatra, West Borneo and Priangan).

These figures do make abundantly clear that variations in the level of within-region inequality are considerable across colonial Indonesia. Even taking into account the fact that these figures are subject to a margin of error (although probably less so than other estimates of inequality in the Global South for before the 1960s, as those generally are not based on such detailed income tax records), it becomes clear that inequality differs drastically between the various regions and that a single statistic about the level of inequality for the entire archipelago provides us with little useful information about levels of inequality and its possible drivers in colonial Indonesia and it makes sense to disaggregate such numbers to smaller geographical units.

⁹ See, for example, the differences in wages shown in Houben (2014), or GDP per capita estimates for the Outer Islands and Java shown in Foldvari et al. (2013: 84).

4 Globalization and Inequality

How we can explain these variations in the levels of inequality? Both scholars who analysed inequality in colonial Indonesia (e.g. Booth 2015, 2019; Leigh and van der Eng 2010; van Leeuwen and Földvári 2017) and those who have investigated the causes of rises and declines in inequality across the globe more broadly (Sokoloff and Engerman 2000; Piketty 2014; Milanovic 2016; Scheidel 2018) have emphasized the role of global trade and (colonial) institutions in driving inequality. Williamson (2011, 2015) employed Heckscher–Ohlin trade theory to suggest that Indonesian inequality increased during the *belle époque* wave of globalization as a result of the specialization in the production of cash crops; as the production of these crops demanded more land than labour, globalization caused the price of land to increase relative to that of labour. A recent study has employed a large set of price data, including from Indonesia, to essentially confirm Williamson’s findings on the relationship between rising globalization and diverging global incomes for the period up to WWI (Francis 2015).¹⁰

In this section, I will examine to what extent patterns of global trade and institutions in the form of the colonial plantation system explain variation in levels of income inequality between residencies. The relationship is investigated in two ways. First, the inequality estimates for all 32 residencies shown in Fig. 4 are combined over the years 1920, 1924 and 1928 to perform a random effects panel analysis in order to examine the *difference* in levels of inequality across the various regions. The data are treated as a panel with 32 regions and three time periods and random effects. Yet it is important to note that one of the underlying elements of the inequality estimates, namely the distribution of land, does not change over the three years because of lacking data. In Appendix 1 it is shown that different land distributions would not lead to significantly different Ginis, so that it is unlikely that it would change the results. Below an additional fixed effects panel analysis is performed using annual data (for 1920–1928) for a sub-sample of 12 residencies (in the Outer Islands) that have very high coverage in the income tax records were collected in order to assess how *changes* in patterns of trade over time are associated with *changes* in inequality.

To investigate the association between global trade and inequality, data on different indicators were gathered from various primary sources (see Appendix 5 for an overview). First is the total value of exports (in guilders) per capita in order to assess the extent of contact with the global market. Second is the amount of planted estate lands as a share of the total land surface in a residency, as it is expected that the

¹⁰ Francis’ (2015) work emphasizes the importance of looking at actual data from the Global South, as it does alter some of important details in the work of Williamson (2011), whose price data were largely based on markets in the West.

mode of production of the export commodities crucially impacted the distribution of the gains from trades, with plantation agriculture likely causing greater inequality than production that is dominated by smallholders.¹¹

To estimate the relationship between inequality and globalization, in this case export agriculture, on different measures of inequality, the following random effects panel model is estimated:

$$INEQ_{j,t} = \alpha + \beta_1 \log(\text{export})_{j,t} + \beta_2 \log(\text{plantation})_{j,t} + X_{j,t}\gamma + Z_j\delta + \tau_t + \varepsilon_{j,t} \quad (3)$$

where j and t index residency and year, respectively; *INEQ* refers to four different measures of inequality (the Gini, Theil, IER and TIR), *export* is the total value of exports per capita, *plantation* is the amount of estate land relative to the total size of a residency, X is a set of controls that differ in the three years and Z a set of time-invariant geographic controls, τ are year dummies, and ε is the error term.

Various controls were included. Milanovic (2018) recently explored factors influencing historical inequality measures for 41 pre-industrial societies in different parts of the globe. He found that more densely populated and less urbanized countries have lower inequality extraction ratios. Therefore, the analysis also includes the percentage of people living in cities with over 5000 inhabitants and the number of people per squared km. It may further be expected that the share of Europeans in the total population affects the distribution of income (as noted by, for example, Booth 1988, 1998; van Zanden and Marks 2012). Geographic controls, like average annual rainfall, slope and altitude of a residency are also included. These variables are the same over the three years, and we are interested in the cross-regional determinants of inequality, so the model does not contain residency fixed effects. Appendix 4 shows the summary statistics for variables included in the analyses.

Table 2 shows the results of a regression analysis of Eq. 3. It is shown that in particular the share of plantations is related to various measures of inequality. The exports per capita variable is not significantly related to the Gini, Theil, and TIR, but only positively related to the IER (significant at the 10% level when including controls). The coefficient on the share of plantation land, on the other hand, is fairly consistent across the different specifications, including and excluding controls. If area of plantations as a share of the total surface in a region is 1 percent higher, this results in about 1.6 higher Gini points, 1.8 higher Theil points, an IER that is 2.1 points higher. There is also a strong positive correlation between plantations and the TIR, but this turns insignificant when also controlling for urbanization, population density and the share of Europeans. That the share of plantation land is strongly related to inequality is not surprising considering the fact that export agriculture is one of the main sources to generate surplus income in late colonial Indonesia. The estate represents a mode of production where most of the surplus accrues to the

¹¹ It is important to note that these statistics on estates are different from the information about the distribution of land. The estate statistics only concern data about commercial estates, while the land distribution information refers to indigenous landholding, often for subsistence production. In fact, the correlation between the land Gini and the share of planted estate land is negative, but statistically insignificant, in all three years 1920, 1924 and 1928. Data available upon request.

Table 2 Globalization and inequality: full sample, 1920, 1924, 1928

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Gini	Gini	Gini	Theil	Theil	Theil	IER	IER	IER	TIR	TIR	TIR
Log export per capita	.791 (.647)		.38 (.473)	.337 (.649)		.066 (.466)	1.953*** (.558)		1.414* (.738)	.797 (.917)		.103 (.634)
Log plantation (%)		1.654*** (.55)	1.605** (.637)		1.968*** (.616)	1.814** (.81)		2.33*** (.927)	2.137** (1.033)		3.944*** (1.063)	1.667 (1.421)
Log income			2.554 (7.82)			-8.162 (8.123)			-23.351*** (8.499)			6.172 (11.653)
Log population density			-1.251 (1.369)			-1.356 (1.856)			-2.809 (2.051)			3.679 (2.664)
Log urbanization (%)			2.97* (1.717)			3.031 (2.438)			-4.083 (3.214)			1.971 (3.395)
Log Europeans (%)			5.086** (2.345)			9.118*** (2.944)			12.511*** (4.182)			15.92*** (3.613)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	93	93	93	93	93	93	92	92	92	93	93	93
Regions	32	32	32	32	32	32	32	32	32	32	32	32
R-squared	.20	.18	.61	.15	.29	.59	.48	.47	.62	.21	.42	.75

Standard errors are in parentheses

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

Table 3 Globalization and inequality: panel. *Sources:* See text and appendix

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gini	Gini	Gini	Gini	Theil	Theil	Theil	Theil
Log exports per capita	2.465*** (.36)	.326 (.19)	.269 (.23)	.27 (.232)	3.532*** (.541)	.048 (.218)	-.247 (.372)	-.245 (.38)
Log plantation (%)	1.939*** (.055)	2.498** (.866)	3.703*** (.378)	3.671*** (.356)	2.311*** (.217)	2.945** (1.218)	4.266*** (.397)	4.173*** (.42)
Log income				-.31 (.471)				-.905 (.618)
Residency FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Obs.	107	107	107	107	107	107	107	107
Regions	12	12	12	12	12	12	12	12
R-squared	.51	.08	.24	.24	.46	.06	.28	.29

Driscoll–Kraay standard errors are in parentheses

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

owner and management of the estate, while the large mass of plantation labourers earns relatively low incomes. Substantial sums of money earned by estates were sent to the Netherlands (see also Gordon 2012), and while some Indonesians higher up the ladder may also have benefitted, they were not a majority. Globalization led to higher levels of inequality, mainly when export production for the global market was organized via estates.

This analysis has thus established a cross-sectional correlation between the share of planted estate land and of inequality as measured by Gini, Theil and IER. But to what extent can *changes* in the levels of inequality be related to *changes* in exports and the amount of estate land over time. In order to address this question, we can analyse the trends in inequality over period 1920–1928 in a sample of 12 residencies that have a high coverage in the income tax records: Bangka, Bengkulu, Biliton, Jambi, Lampung, Moluccas, Palembang, Riouw, Southeast Borneo, East Sumatra, West Sumatra, and Tapanuli. Data on inequality were gathered for these residencies for each of the eight years for which income tax data exist and related to annual data on exports and estates. These could then be used in a panel analysis, using the following model specification:

$$INEQ_{j,t} = \alpha + \beta_1 \log(\text{export})_{j,t} + \beta_2 \log(\text{plantation})_{j,t} + \beta_2 \log(\text{income})_{j,t} + \delta_j + \tau_t + \varepsilon_{j,t} \quad (4)$$

where the same applies as in the case of Eq. 3, with the main difference being that besides income as the only control variable (due to data availability), the model includes both year (τ .) and residency (δ) fixed effects. There are no annual demographic data available, so it is not possible for to control for urbanization and population densities, but these are unlikely to change much over such a short period

and should thus be captured largely by the residency fixed effects. Similarly, lacking detailed annual price and wage data for these regions, the TIR and IER could not be meaningfully calculated. Table 3 reports the results of the panel regressions. Columns 1–4 report the results for the Gini as a measure of inequality, while Columns 5–8 report the results with the Theil index as inequality measure. Columns 1 and 5 report between region results (without time and regional fixed effects). In Columns 2 and 6 regional fixed effects are added and in Columns 3 and 7 year fixed effects. Finally, Columns 4 and 8 also include average income as a control variable. S I estimate Eq. 4 using Driscoll–Kraay standard errors to account for serial and spatial correlation (Driscoll and Kraay 1998).

Looking at the cross-sectional difference in levels between these twelve regions over the entire period, without any controls, both exports per capita and the share of plantation land are positively associated with levels of inequality in these 12 residencies, which is also more or less what we have seen for the full sample. Once we include regional (Columns 2 and 6) and time fixed effects (Columns 3 and 7), however, the association between exports per capita and inequality disappears, while the coefficient on the plantation variable increases. Changes in average income have no effect on these results (Columns 4 and 8). If the share of plantations increases by 1 percent, this results in the increase of the Gini of 3.7 Gini points and some 4.2 Theil points. The relationship between plantations and inequality suggests the importance of the institutional context in which global trade is taking place.

The analyses based on Eqs. 3 and 4 suggest that global trade indeed affected inequality levels and trends. The measure on the total value of exports per capita is cross-sectionally positively correlated with the Gini and the Inequality Extraction Rate for the full sample. It thus explains part of the difference in inequality between regions with high levels of exports and those regions less integrated in the world economy. Yet, exports per capita do a poor job in explaining trends in inequality over time, as shown by the panel analysis. Much more consistent both in the cross-sectional analysis and in the panel is the effect of the area planted by estates. This affects all measures of inequality in all specifications and the effect is sizeable. In Appendices 2 and 3 it is shown that this result holds when using a wide range of assumptions in computing inequality, which gives a high degree of confidence in the results.

Thus far, we have examined correlations between the extent of trade and planted estate area, and inequality, but it remains unclear in what direction the relationship runs. It may be the case that estates are set up in regions that are characterized by high inequality, in order to benefit, for example, from low wages. It is outside the scope of this paper to tease out the precise causal mechanisms underlying this relationship, yet it is worthwhile to provide some further information about the origins and nature of plantation agriculture in the Dutch East Indies by examining local and colonial institutions pertaining to land.

Privately owned European plantation agriculture in the Dutch East Indies started only in 1870, with the implementation of the Agricultural Law. Before that, all export cultivation was monopolized by the Dutch colonial government. Under the Agricultural Law, irrigated *sawah* lands were the property of the local population and could not be owned by European enterprises. These lands could only be rented

by Europeans on a short-term basis from the local population. At the same time, the Agrarian Law, and the accompanying Domain Declarations in the 1870s, declared all still unused “waste lands” to be the property of the colonial government. European enterprises could obtain long-term leases (for 75 years) on these lands for the price of 1 guilder per *bouw* annually (Djalins 2012: 97). Most plantations in the East Indies, with the exception of sugar plantations in Java that required irrigated fields, were established on these waste lands under these conditions. While in theory, the Dutch colonial government thus became owner of all unused lands in the early 1870s, in practice, the Dutch had to take into account local customary law (*adat*) with regard to land rights, which differed substantially from region to region. Whereas in some regions, villages and their inhabitants had relatively strong user rights to nearby waste lands, in others, these waste lands were considered the property of the sovereign or local aristocracy. Secondary literature suggests that the leasing of these waste lands to estates often had to be negotiated with local elites and that the colonial government feared social uprisings if it neglected local *adat* (e.g. von Benda-Beckmann and von Benda-Beckmann 2013; de Zwart 2020).

Additionally, local geographical conditions may have played a role in the location of plantations. Following Sokoloff and Engerman (2000) it may be expected that the suitability for growth of specific crops impacted the share of planted estate land. They argue that some commodities, like sugar and tobacco, have greater returns to scale and therefore favoured production via plantations. Rubber and coffee, on the other hand, can also be efficiently produced by smallholders (Nugent and Robinson 2010; Ross 2017). Additional factors that may have influenced the share of plantations include the distance to the colonial centre, Batavia, or the total years of colonial occupation. More research into the determinants of plantation land is necessary, but it does seem clear that plantations are determined by a variety of variables exogenous to the models presented here.

5 Conclusions

In this paper, levels of inequality for Indonesia were reconstructed for 32 residencies in the late colonial period. Because inequality estimates are based on a variety of assumptions, four different measures of inequality were computed: Gini, Theil, Inequality Extraction Rate and the Top Income Rate. These different measures show a strong correlation which gives confidence in the results. It was shown that levels of inequality differed radically across the archipelago. Poor regions without any export activity, like Banten and Madura, had very low Gini coefficients in the 1920s. Slightly richer regions with export activity dominated by indigenous smallholders, such as West Sumatra and Jambi, had Ginis in the 1930s. Areas with higher incomes and more commercial activities that were dominated by Europeans, like Batavia, Surabaya and Semarang, had very high Gini coefficients of above 50. While these calculations may to some extent be influenced by differences in levels of enforcement, it seems unlikely that the geographical distribution of inequality has radically altered because of this. The data clearly suggest that even within a single colony, levels of inequality varied greatly and this puts some doubts on the representativeness

of using a single number to capture the level of inequality in a large economy: as regions with very low levels of inequality can exist in countries with high overall inequality and vice versa. This paper thus emphasizes the importance of moving beyond the country level to more detailed regional analyses, in particular if one wants to understand the factors that underlie levels and trends in inequality.

Secondly, it was examined to what extent global trade may have contributed to inequality, as this has featured prominently in the literature on the determinants. It is shown that in particular the total share of land that was being used for plantation agriculture in a residency is related to various measures of inequality. Additionally, a panel analysis including 8 years of data for a sub-sample of twelve residencies which had a high coverage in the income tax records suggests that increases in plantation land is also strongly related to increases in inequality. Thus, what matters for inequality is not only the rise of global trade per se, but more so how this trade was organized: via plantations or via smallholders. These analyses were based on a very limited number of observations (there were only 32 residencies with good enough data), referring to only one country, so future research for other areas and time periods is needed to see whether the same patterns can be observed. The fact that there is a substantial degree of consistency in the outcomes of the regressions, as shown in both the main paper and the appendices, gives confidence in these results. The results confirm the (largely theoretical) suggestions by, for example, Williamson (2011) and Sokoloff and Engerman (2000) on the consequences of trade in tropical commodity exports and plantations for inequality. The plantation is an institution that we associate with high inequality with most of the surplus accruing to the owners, while the mass of labourers remains relatively poor. In this colonial context, high inequality in some of these regions may well have impeded long-run economic development (Halter et al. 2014), but more research is needed to investigate this relationship for Indonesia. Further research also needs to be done on what determines the spread of plantations in the Dutch East Indies, as well as within other countries in the Global South more generally.

Appendices

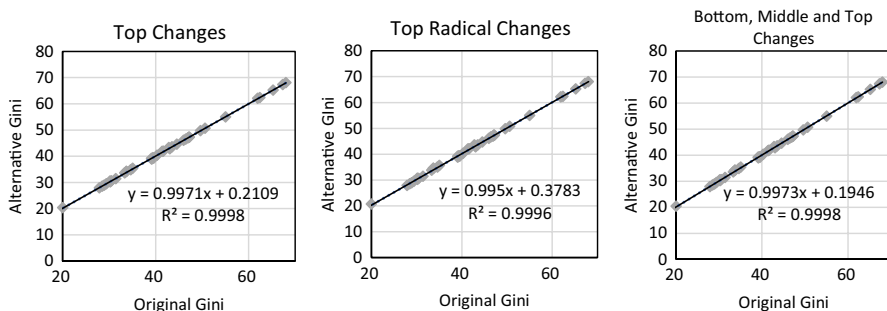
Changing land inequality

No data on the full distribution of land in the different residencies of Java are available for the 1920s. In the main text, use is made of the distribution of land as reported in the Declining Welfare Study from 1903 (Hasselmann 1914: Appendix R). It is likely that in the period between 1903 and 1920, as well as from 1920 to 1928, land inequality increased (e.g. Scheltema 1927). The report by Meijer Ranneft and Huender (1926) also noted further shifts in patterns of landholding. They did not give data about the entire distribution of plots, as was available for 1903, but they did note how the number of landowners with over 25 *bouw* of land increased, reproduced in Table 4.

The table makes clear that, with the exception of Rembang, Surabaya and Pekalongan, there were increases in the number of larger landowners. In order to see how

Table 4 Number of large landholding households in Java, 1903 versus 1925. *Sources:* Hasselman (1914: Appendix R) and Meijer Ranneft and Huender (1926: Appendix VIII)

	No. of households with over 25 <i>bouw</i> land in 1903	No. of households with over 25 <i>bouw</i> land in 1925	Difference
Banten	12	157	145
Batavia	13	376	363
Preanger	1028	1226	198
Cirebon	38	268	230
Pekalongan	334	106	− 228
Semarang	164	250	86
Banyumas	47	207	160
Rembang	43	43	0
Surabaya	160	79	− 81
Kedu	36	80	44
Madiun	66	78	12
Madura	32	50	18
Kediri	99	107	8
Pasuruan	77	137	60
Besuki	25	223	198

**Fig. 6** Correlation between computed Gini in main text and different scenarios of changes in the land distribution

this would impact the overall distribution, it may be assumed that this concentration led to a greater degree of landlessness, or was mainly the result of the acquisition of newly cleared lands (1). In this case, the additional large landowning households (column “Difference” in Table 4) can be added to the distribution from 1903. In doing so, I use similar shares of landowners of 25, 30, 35 and over 35 *bouw* to distribute the added landowning households over the original information on landownership. In a more radical—and more unlikely—scenario, all the additional landowners are added to the largest class of landowners (those with 35 or more *bouw*) (2). Third, it may also have been the case that this concentration of land went at the expense of patches of land from landowners in the middle of the distribution, and

that an added number of households became small landowners (plots < 1 *bouw*) (3). This is less likely, given what we know about changes in landholding patterns as these more often went at the expense of small landowners, but making this assumption, we can test how an upper-bound assumption of increases in land inequality would affect the overall computed Ginis. To compute this scenario of larger changes in inequality, similar numbers of additional households (as in Column “Difference” in Table 4) are assigned small plots (up to 1 *bouw*), while a group of middle-class landowners (with plots between 2 and 20 *bouw*) is reduced by double the number of households from Column “Difference” in Table 4.

Figure 6 shows Ginis for Java in 1920, 1924 and 1928 as shown in the main text, in addition to scenarios 1–3 mentioned above. It becomes immediately clear from these figures that these changes, even in more unlikely scenarios of changes in land inequality up to the 1920s, the calculated Ginis are not affected as the correlation is almost perfect and the coefficient is 0.99. It is thus unlikely that the possible changes in the land distribution between 1903 and 1925 will influence the results.

Wage and price data

Rural and urban unskilled wages for all residencies (both Java and Outer Islands) as paid by the Public Works Department are available for the years 1920 to 1924 (Malines van Ginkel 1926, vol. I: 207–219). Rural wages were used in this paper. Additionally, for those residencies that produced sugar (i.e. Cirebon, Pekalongan, Banyumas, Kedu, Madiun, Surabaya, Kediri & Pasuruan), wages paid for unskilled workers in the sugar industry were available up to 1940. On the basis of the correlation between wages in the sugar industry and wages paid for by the Public Works Department, the wages of the latter were extrapolated to 1928 (data files showing these correlations for all the residencies separately are available upon request). For other residencies without a sugar industry, like Banten, Batavia and Priangan, correlations of wages with neighbouring provinces were used to extrapolate to 1928. Because these wages are very stable over time (the average coefficient of variation taken over the years 1920–1929 for all residencies is 0.08), these procedures would hardly affect the results. For the Outer Islands, wages were used to calculate TIR. For these residencies, wage data were only available for the period 1920–1924 and there was no sugar industry in the Outer Islands. Because wages were so stable over these years (c.v. of 0.07), wage levels of 1924 were extrapolated to 1928.

Price data on the basic articles of consumption shown in Table 1 were gathered from two original publications (CKS [Centraal Kantoor voor de Statistiek] 1927, 1931). For the rice price, the cheapest variety was taken (*bras merah*) (CKS [Centraal Kantoor voor de Statistiek] 1927, 1931), prices were reported per *gantang*, equal to 8.58 litre or 6.7 kg. Similarly, for soya beans, cassava and sweet potatoes, prices for the cheapest variety were collected. These prices were reported per *picul*, a measure that is equal to 61.76 kg. Dried fish and sugar were reported per *cat-tie*, which was 0.62 kg. For lamp oil the price of coconut oil was taken. This was reported per bottle and it was assumed that one bottle equalled 1 litre. Data on salt

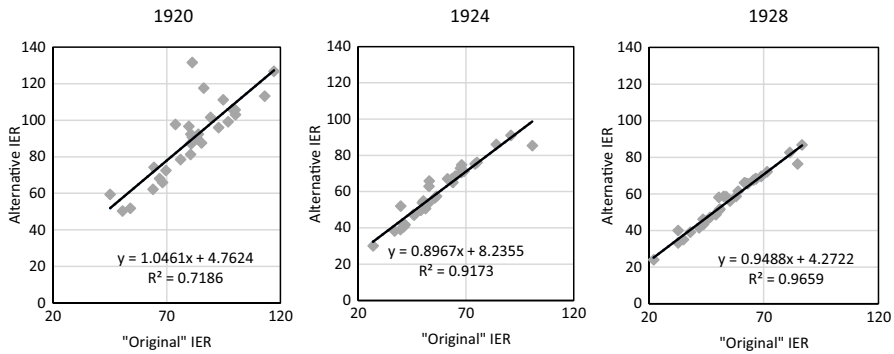


Fig. 7 Correlation between original and alternative IER, 1920, 1924 and 1928

and cotton were obtained through personal correspondence with Jan Luiten van Zanden, who employed these data in the publication van Zanden and Marks (2012). Datasheets are available upon request. Fuel (petroleum) price data were obtained from Korthals Altes (1994: 133–134). Petroleum litre prices were converted into MBTU using a conversion sheet from Iowa State University (2018).

On the basis of these sources, it was possible to compute the cost of subsistence for Batavia (present-day Jakarta). Only for rice in 1928, there is data available on prices in all different residencies. In order to get to a different basket across the archipelago, the price differential in rice between each province and Batavia was assumed to be constant across years and products. This implies that if, for example, rice was 10 percent more expensive in East Sumatra than in Batavia in 1928, the basket price for East Sumatra was 1.1 times the basket price in Batavia for all three years. This was the procedure followed in the main text. It could also be the case that because rice is a cheap bulk good, relatively expensive to transport, the price differentials in rice actually present an upper bound in overall price differentials. Possibly, the other goods did not show a similarly large price difference. To test how this assumption affects the results, the price differential in rice prices is only used to calculate the total cost for rice, and all the other products are assumed to be similarly priced as in Batavia. Figure 7 shows the correlation between the IER as calculated in the main text and the alternative assumption of similar prices as in Batavia for all non-rice goods and only different rice prices. With the exception of a few outliers in 1920, the correlation between the two measures is quite strong and the coefficients suggest relatively small differences of between 4 (average increase in 1920), 5 and 11 percent (decrease in 1928 and 1924, respectively).

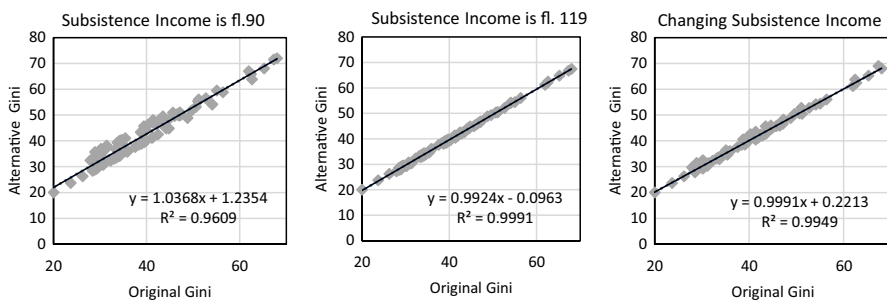
Furthermore, we can test to what extent the results as obtained in the main text are robust to these adjustments. Table 5 shows the results of running Eq. 3 with both the data used in the main text and the alternatively calculated IER. The coefficients change slightly and the errors increase a bit. The coefficient on

Table 5 Robustness test: full panel with alternative IER. *Sources:* See text and appendix

	(1)	(2)	(3)	(4)	(5)	(6)
	Original	Original	Original	Alternative	Alternative	Alternative
Log export per capita	1.953*** (.558)		1.414* (.738)	3.363*** (.64)		2.991*** (.94)
Log plantation (%)		2.33** (.927)	2.137** (1.033)		2.859*** (1.033)	2.035* (1.114)
Log income			− 23.351*** (8.499)			− 25.439*** (9.495)
Log population density			− 2.809 (2.051)			− 1.403 (1.944)
Log urbanization (%)			− 4.083 (3.214)			− .971 (3.348)
Log Europeans (%)			12.511*** (4.182)			8.605** (4.387)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	92	92	92	92	92	92
Regions	32	32	32	32	32	32
R-squared	.48	.47	.62	.60	.56	.69

Standard errors are in parentheses

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

**Fig. 8** Correlation between original and alternative Ginis. *Sources:* See text and appendix

plantation share decreases and the error increases, so that statistical significance level drops. The coefficient on exports per capita on the other hand, increases in size and statistical significance.

A.3. Subsistence incomes

As noted in the main text, households with incomes below fl. 120 were exempted from the income tax. As it is unclear what the average income was, these households

have been assigned fl. 115 per annum, as fl. 120 was already considered not much above subsistence level. Here, three different correlations are shown to suggest what different assumptions on the average income of this subsistence group do. In the first scenario, we use information on the cost of subsistence from Table 1 in the main text, which shows that this was fl. 150 in 1920, fl. 105 in 1924 and fl. 90 in 1928. I replace fl. 150 with fl. 119, as households earning fl. 150 would be included in the income tax. In this scenario, subsistence income changes across these different years. In a second scenario, the fl. 90 is used for all three years as a lower bound estimate, while in a third scenario fl. 119 is used as annual subsistence income in all three years as an upper-bound estimate. Figure 8 shows the correlations between these three new estimates and the estimate as in the main text using fl. 115 as “subsistence income” all three years. From the figure it becomes clear that these different assumptions hardly affect the Gini calculated. Only in the most radical scenario, where average income of the “subsistence” households is fl. 90 per annum, which, in most of these years would mean a standard of living below the barebones subsistence level, does the Gini increase slightly, with about 4–5 Gini points on average. The very strong correlation (R^2 of between 0.96 and 0.99) suggests that using these figures, rather than the “original” used in the main text, in a regression analysis, would not affect the outcomes of this article.

To test whether regressions results of Table 2 in the main text hold, Table 6 shows the results of the same regression with the variation in assumptions regarding subsistence incomes. The coefficients for the effect of exports and plantation shares on the Ginis calculated with subsistence incomes of fl. 90 differ slightly, but not radically (shown in Columns 1–3). The results in Columns 4–9 of Table 6 are really almost identical to those in Table 2 (as one would expect on the basis of Fig. 8). The results shown in the main text hold.

Additionally, it was also tested whether the results of the analysis on the association between plantations and Gini would hold when using the “Taxpayer Gini” without any adjustments (see the correlation between the two measures in the main text, Fig. 1). Table 7 shows both the analysis as in Table 2 (Columns 1–3) and the same analysis with the “Taxpayer Gini” (Columns 4–6). However, the coefficients are somewhat different, especially in the analyses without any controls. The results of the models run with all the control variables included are remarkably similar with regard to the exports per capita and plantation variables. Some controls, however, show radically different coefficients and signs, but these are not of our concern in this analysis. The results in the main text hold, and in fact, become stronger in the regressions using the taxpayer Gini estimates.

Table 6 Robustness test, full sample panel with alternative Ginis. *Sources:* See text and appendix

	(1) Gini90	(2) Gini90	(3) Gini90	(4) Gini119	(5) Gini119	(6) Gini119	(7) Changing Gini	(8) Changing Gini	(9) Changing Gini
Log export per capita	.84 (.661)		.557 (.536)	.784 (.645)		.353 (.463)	.869 (.708)		.436 (.499)
Log plantation (%)		1.824*** (.572)	1.64** (.694)		1.632*** (.549)	1.608** (.625)		1.651*** (.54)	1.591** (.633)
Log income			.022 (6.526)			2.893 (7.996)			2.397 (7.893)
Log population density			-.869 (1.423)			-1.322 (1.361)			-1.179 (1.366)
Log urbanization (%)			3.576* (2.078)			2.858* (1.669)			3.239* (1.658)
Log Europeans (%)			4.847* (2.615)			5.098** (2.307)			4.986** (2.379)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	93	93	93	93	93	93	93	93	93
Regions	32	32	32	32	32	32	32	32	32
R-squared	.23	.20	.55	.19	.18	.62	.20	.18	.60

Standard errors are in parentheses

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

Table 7 Robustness test, full sample panel with taxpayer Ginis. *Sources:* See text

	(1)	(2)	(3)	(4)	(5)	(6)
	Corrected Gini	Corrected Gini	Corrected Gini	Taxpayer Gini	Taxpayer Gini	Taxpayer Gini
Log export per capita	.791 (.647)		.38 (.473)	.163 (.416)		.489 (.376)
Log plantation (%)		1.654*** (.55)	1.605** (.637)		3.743*** (.553)	1.855*** (.563)
Log income			2.554 (7.82)			-4.227 (3.798)
Log population density			-1.251 (1.369)			4.356*** (.932)
Log urbanization (%)			2.97* (1.717)			2.099 (1.75)
Log Europeans (%)			5.086** (2.345)			4.68** (2.093)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	93	93	93	93	93	93
Regions	32	32	32	32	32	32
R ²	.19	.18	.61	.11	.64	.80

Standard errors are in parenthesis

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

Summary statistics and missing data

Table 8 shows the summary statistics of the data included in the cross-sectional analysis. In principle, three years of data for 32 regions should result in a total of 96 observations for each variable. However, for a couple of years, missing observations or erroneous outliers led to a reduction of observations. Missing data are:

- Gini, Theil, TIR and average income: too few people were paying the income tax in Bali in 1920. The combined data on income from land for Rembang in 1928 suggested the larger ones plot the less income one would generate (due to low land productivity and relatively high wages), and this was considered unlikely and therefore omitted from this paper.
- IER: the same applies as in the case of the Gini and Theil, only one additional observation is lost, as the calculations for Aceh in 1920 resulted in an implausible IER of 585. This was considered an erroneous outlier (nine standard deviations from the mean) and removed from the analysis.

Not included at all in this article are the residencies of Yogyakarta, Surakarta, Timor, Ternate and West New Guinea, as the data were insufficient to compute any inequality measure.

One further observation is lost when transforming variables into logs as there are no plantations in Banka in 1920. All regressions in the text and appendix are run without Bali 1920, Banka 1920 and Rembang 1928. Regressions with IER as outcome furthermore are without Aceh 1920.

Table 9 shows the summary statistics for the panel analysis with the sub-sample of 12 residencies. Similar to the cross section above, Banka 1920 is lost due to the log transformation of the plantation variable.

Table 8 Summary statistics of cross-sectional data. *Sources:* See Appendix 5

Variable	Regions	Obs.	Mean	SD	Min	Max
Gini	32	93	38.15	11.42	13.34	68.03
Theil	32	93	24.67	14.30	4.83	66.96
TIR	32	93	36.76	24.30	8.54	137.56
IER	32	92	64.14	19.96	22.08	117.14
Avg. income	32	93	296.19	99.29	143.35	772.25
Plantation (%)	32	93	2.02	2.71	0.0039	9.60
Exports/capita	32	93	53.28	86.62	0.002	507.5
European (%)	32	93	0.34	0.317	0.02	1.49
Population (km ²)	32	93	150.86	154.49	2.63	447.14
Urbanization rate	32	93	5.36	2.93	1.17	15.84
Altitude (m)	32	93	304.08	190.76	27.19	649.17
Rain (mm)	32	93	2516.5	436.83	1615.74	3306.56
Slope (%)	32	93	2.57	1.48	0.44	5.97

Table 9 Summary statistics for panel data. *Sources:* See text and Appendix 5

Variable	Regions	Obs.	Mean	SD	Min	Max
Gini	12	107	36.58	7.25	23.75	56.86
Theil	12	107	25.23	10.22	8.37	56.53
Avg. income	12	107	378.62	114.99	17.93	972.13
Exports/capita	12	107	94.31	125.79	0.27	555.26
Plantation (%)	12	107	0.36	0.68	0.0039	3.27

Detailed data sources

To calculate income from land across different residencies in Java:

- Land distribution across Java in 1903: Hasselman (1914: Appendix R).
- Paddy production in tons per hectare in 1920, 1924 and 1928: Boomgaard and Van Zanden (1990: 96, 121).
- Paddy price: only for 1928 data on the market price for paddy were available for all residencies across Java (SJNI 1928: 290–291). The price differentials across the different residencies in 1928 were then assumed to be similar also in 1920 and 1924 and the Batavia paddy price (from Van Zanden and Marks 2020) for those years was used to extrapolate the prices for the other residencies.
- Wages: Malines van Ginkel (1926, vol. I: 207–219) (for details see discussion in Appendix 2).
- Labour participation rate in 1930: to get an impression of how many people were working per household and how this differed across Java, the total number of people with an occupation were divided by the total population, using data from EZ [Dept. van Economische Zaken] (1936: 126–127).

Variables for regression analyses:

The main text already describes the sources and calculations for the inequality measures; the focus here is therefore on the other variables.

Population data

The combined data on the indigenous, European, Chinese and other Asian population were used to compute total population in 1920 using data from Boomgaard and Gooszen (1991). For 1924 and 1928, there were observations for the indigenous population, but for the European, Chinese and other Asian population data were only available for 1920 and 1930. These figures were linearly interpolated and combined with the series on the indigenous population. These population figures were divided by the total surface of each residency as reported in CBS [Centraal Bureau voor de Statistiek] (1922).

Table 10 Mergers of residencies in 1928. *Source:* Cribb (2000)

Residency	Merger of
Batavia	Batavia, Krawang and Buitenzorg
Cirebon	Cirebon and Indramayu
Priangan	West, East and Middle Priangan
Pekalongan	Pekalongan and Tegal
Semarang	Semarang, Kudus and Blora
Rembang	Rembang and Bodjonegoro
Surabaya	Grissee, Surabaya and Modjokerto
Bayumas	North and South Wonoboso
Kedu	Kedu and Bagelen
Madiun	Madiun and Ponorogo
Kediri	Kediri and Blitar
Pasuruan	Pasuruan, Malang and Probolinggo
Besuki	Djember and Bondowoso

Urbanization

To compute the rate of urbanization, the number of people in towns over 5000 people was taken from Boomgaard and Gooszen (1991) for 1920 and 1930. These numbers were linearly interpolated to arrive at figures for 1924 and 1928 and then divided by the total population in each residency as calculated above.

Exports

Total value of exports per person: Computing exports per capita is straightforward for residencies in the “Outer Islands” as these figures can be taken from the research of Touwen (2001). To get at the total export of Riouw in this paper, figures for Riouw free-trade zone and Indragiri were combined. For Java, no breakdown per residency of exports is possible, as the exports from Java are enumerated by main ports of Java: because some ports serve multiple residencies, and some residencies are served by multiple ports, it is virtually impossible to get from these figures to an accurate number for the residencies. Instead, we have to rely on production as enumerated in LNH [Dept. van Landbouw, Nijverheid en Handel] (1922, 1924, 1925a, b; 1928b, 1929a) and CBS [Centraal Bureau voor de Statistiek] (1922). These publications show production of estates as well as provide estimates of production by native producers and give figures on petroleum exports per residency. The volumes are converted into total values in guilders by taking the export price of the various products at Batavia from Korthals Altes (1994) and LNH (1925a, b). For the panel analysis, data from Touwen (2001) sufficed.

Plantations

In order to estimate share of plantations, figures on the planted area of estate land from the LNH (1922, 1925b, 1929a, b) were taken. For the panel analysis, all consecutive years of this same source were employed. These figures on planted estate area were then divided by the total surface area in each residency as reported in CBS [Centraal Bureau voor de Statistiek] (1922).

Geographic controls

Geographical control variables consist of (1) the average rainfall per annum in the period 1970–2000 (Fick and Hijmans 2017) and (2) average slope and (3) average altitude from (CSI and CGIAR 2019).

Changing residency borders

Residency borders were changed in Java in 1925 as many new residencies were created out of the existing ones (see Cribb 2000). In order to look at consistent residencies over time, data for the separate residencies were merged in order to get the same residencies as in 1920 and 1924. Data for the following residencies were merged for 1928 (Table 10).

Acknowledgements I thank Sylvia Dooper and Phylcia Soekhradj for valuable research assistance and Pim Arendsen for help with creating the maps and the collection of geographic data. I gratefully acknowledge financial support from the Netherlands Organisation for Scientific Research (NWO) for the project “Unfair Trade? Globalization, Institutions and Inequality in Southeast Asia, 1830–1940” (NWO Veni Grant No. 275-53-016). Seminar participants at the International Institute for Social History, Amsterdam, Wageningen University and Utrecht University provided useful comments and suggestions. I thank Ewout Frankema, Michiel de Haas and Daniel Gallardo Albarran, Tanik Joshapura as well as the anonymous referees and editor of this journal, for useful remarks and suggestions. Any remaining errors and dubious interpretations are mine alone.

Data availability Underlying data and stata code for replication are accessible via <https://doi.org/10.3886/E130624V1>.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Official Sources

- CBS [Centraal Bureau voor de Statistiek] (1922) Jaarcijfers voor het Koninkrijk der Nederlanden: Koloniën 1920. [Yearbook for the Kingdom of the Netherlands: colonies 1920]. Belinfante, The Hague
- CKS [Centraal Kantoor voor de Statistiek] (1927) Prijzen, indexcijfers en wisselkoersen op Java, 1913–1926. Mededeelingen van het Centraal Kantoor voor de Statistiek No. 46. Landsdrukkerij, Batavia
- CKS [Centraal Kantoor voor de Statistiek] (1931) Prijzen, indexcijfers en wisselkoersen op Java, 1913–1929. Mededeelingen van het Centraal Kantoor voor de Statistiek No. 88. Landsdrukkerij, Batavia
- EZ [Dept. van Economische Zaken] (1936) Volkstelling 1930. Vol. VIII Overzicht voor Nederlandsch-Indië. [Census of 1930 in the Netherlands Indies. Vol. VIII (Summary)]. Landsdrukkerij, Batavia
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1922) Beplante uitgestrektheden en producties in het Grootlandbouwbedrijf in Nederlandsch Indië in de jaren 1920 en 1921. Mededeelingen van het Statistisch Kantoor, No. 7. Landsdrukkerij, Weltevreden
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1924) Statistisch Jaaroverzicht van Nederlandsch-Indië 1924. Weltevreden
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1925a) Jaaroverzicht van den in- en uitvoer van Nederlandsch-Indië gedurende het jaar 1924. Mededeelingen van het Centraal Kantoor voor de Statistiek No. 20. Landsdrukkerij, Weltevreden
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1925b) Beplante uitgestrektheden en producties in het grootlandbouwbedrijf in Nederlandsch Indië in 1924. Mededeelingen van het Centraal Kantoor voor de Statistiek No. 22. Landsdrukkerij, Weltevreden
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1928a) Korte berichten voor landbouw, nijverheid en handel. Departement van Landbouw, Buitenzorg
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1928b) Statistisch Jaaroverzicht van Nederlandsch-Indië 1928. Weltevreden
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1929a) De Landbouwexportgewassen van Nederlandsch-Indië in 1928. Beplante uitgestrektheden, producties en uitvoeren. Mededeelingen van het Centraal Kantoor voor de Statistiek No. 74. Landsdrukkerij, Weltevreden
- LNH [Dept. van Landbouw, Nijverheid en Handel] (1929b) Jaaroverzicht van den in- en uitvoer van Nederlandsch-Indië gedurende het jaar 1928. Mededeelingen van het Centraal Kantoor voor de Statistiek No. 79. Landsdrukkerij, Weltevreden
- KV (1922) Koloniaal verslag over 1922. Bijlagen van het Verslag der Handelingen van de Tweede Kamer der Staten Generaal. Algemeene Landsdrukkerij, The Hague
- KV (1926) Verslag van Bestuur en Staat van Nederlandsch-Indië, Suriname en Curaçao over 1926. Bijlagen van het Verslag der Handelingen van de Tweede Kamer der Staten Generaal. Algemeene Landsdrukkerij, The Hague
- KV (1930) Verslag van Bestuur en Staat van Nederlandsch-Indië, Suriname en Curaçao over 1930. Bijlagen van het Verslag der Handelingen van de Tweede Kamer der Staten Generaal
- Staatsblad (1920) Staatsblad van Nederlands-Indië over het jaar 1920. Landsdrukkerij, Weltevreden
- Staatsblad (1921) Staatsblad van Nederlands-Indië over het jaar 1921. Landsdrukkerij, Weltevreden

Literature

- Allen RC (2015) The high wage economy and the industrial revolution: a restatement. *Econ Hist Rev* 68(1):1–22
- Allen RC (2017) Absolute poverty: when necessity displaces desire. *Am Econ Rev* 107(12):3690–3721
- Allen RC, Bassino J-P, Ma D, Moll-Murata C, Van Zanden JL (2011) Wages, prices, and living standards in China, 1738–1925: in comparison with Europe, Japan, and India. *Econ Hist Rev* 64(February):8–38
- Apergis N, Dincer O, Payne JE (2013) Economic freedom and income inequality revisited: evidence from a panel error correction model. *Contemp Econ Policy* 32:67–74

- Bennett DL, Nikolaev B (2017) On the ambiguous economic freedom-inequality relationship. *Empir Econ* 53(2):717–754
- Boeke JH (1955) *Economie van Indonesie*. H.D. Tjeenk Willink & Zoon, Haarlem
- Boomgaard P, Gooszen AJ (1991) *Changing economy in Indonesia 11: population trends 1795–1942*. KIT Press, Amsterdam
- Boomgaard P, Van Zanden JL (1990) *Changing economy in Indonesia 10: foodcrops and arable lands, Java 1815–1942*. KIT Press, Amsterdam
- Booth A (1980) The burden of taxation in colonial Indonesia. *J Southeast Asian Stud* 11(1):91–109
- Booth A (1988) Living standards and the distribution of income in colonial Indonesia: a review of the evidence. *J Southeast Asian Stud* 19(2):310–334
- Booth A (1998) *The Indonesian economy in the nineteenth and twentieth centuries. A history of missed opportunities*. Macmillan, Basingstoke
- Booth A (2015) A century of growth, crisis, war and recovery, 1870–1970. In: Coxhead I (ed) *Routledge handbook of Southeast Asian economics*. Routledge, London and New York, pp 43–59
- Booth A (2019) *Living standards in Southeast Asia. Changes over the long twentieth century, 1900–2015*. Amsterdam University Press, Amsterdam
- Bosma U (2019) *The making of a periphery how Island Southeast Asia became a mass exporter of labor*. Columbia University Press, New York
- Breman J (1989) *Taming the coolie beast: plantation society and the colonial order in Southeast Asia*. Oxford University Press, Oxford
- Breman J (2010) *Koloniaal Profijt van Onvrije Arbeid. Het Preanger Stelsel van Gedwongen Koffieteel Op Java*. Amsterdam University Press, Amsterdam
- Cribb R (2000) *Historical Atlas of Indonesia*. University of Hawaii Press, Honolulu
- CSI, CGIAR (2019) SRTM 90 m Digital Elevation Database 4.1. 2019. <https://cgiasi.community/data/srtm-90m-digital-elevation-database-v4-1/>
- de Zwart P (2020) Globalisation, inequality and institutions in West Sumatra and West Java, 1800–1940. *J Contemp Asia*. <https://doi.org/10.1080/00472336.2020.1765189>
- Djalins UWM (2012) *Subjects, lawmaking and land rights: agrarian regime and state formation in late-colonial Netherlands East Indies*. Cornell University, Ithaca
- Driscoll JC, Kraay AC (1998) Consistent covariance matrix estimation with spatially dependent panel data. *Rev Econ Stat* 80(4):549–559
- Dros N (1992) *Changing economy of Indonesia 13: wages 1820–1940*. KIT Press, Amsterdam
- Elson RE (1994) *Village Java under the cultivation system, 1830–1870*. Allen & Unwin, Sydney
- Fick SJ, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37(12):4302–4315
- Foldvari P, Van Leeuwen B, Marks D, Gall J (2013) Indonesian regional welfare development, 1900–1990: new anthropometric evidence. *Econ Hum Biol* 11(1):78–89
- Francis JA (2015) The periphery's terms of trade in the nineteenth century: a methodological problem revisited. *Hist Methods J Quant Interdiscip Hist* 48(1):52–65
- Frankema E, Marks D (2010) Growth, stability, but What about equity? Reassessing Indonesian Inequality from a comparative perspective. *Econ Hist Dev Regions* 25(1):75–104
- Geertz C (1963) *Agricultural involution. The processes of ecological change in Indonesia*. University of California Press, Berkeley, Los Angeles and London
- Geloso V (2018) The fall and rise of inequality: disaggregating narratives. *Adv Austrian Econ* 23:161–175
- Geloso V, Magness P (2020) The great overestimation: tax data and inequality measurements in the United States, 1913–1943. *Econ Inq* 58(2):834–855
- Gordon A (2012) How big was Indonesia's 'Real' colonial surplus in 1878–1941? *J Contemp Asia* 42(4):560–580
- Halter D, Oechslin M, Zweimüller J (2014) Inequality and growth: the neglected time dimension. *J Econ Growth* 19(1):81–104
- Hasselman CJ (1914) *Algemeen Overzicht van de Uitkomsten van Het Welvaart-Onderzoek Gehouden Op Java an Madoera in 1904–1905*. Nijhoff, The Hague
- Holcombe RG, Boudreaux CJ (2016) Market institutions and income inequality. *J Inst Econ* 12(2):263–276
- Houben V (2014) *Wages, living standards and urban–rural inequality in the outer islands of colonial Indonesia 1913–1924*

- Iowa State University (2018) Liquid fuel measurements and conversions. 2018. <https://www.extension.iastate.edu/agdm/wholefarm/pdf/c6-87.pdf>
- Kian-wie T (1969) Plantation agriculture and export growth: an economic history of East Sumatra, 1863–1942. University of Wisconsin Press, Madison
- Korthals Altes WL (1994) Changing economy in Indonesia, vol. 15: prices (non-rice) 1814–1940. KIT Press, Amsterdam
- Leigh A, van der Eng P (2010) Top incomes in Indonesia, 1920–2004. In: Atkinson AB, Thomas P (eds) Top incomes over the twentieth century: a global perspective. Oxford University Press, Oxford, pp 171–219
- Lindert PG, Williamson JG (2016) Unequal gains: American growth and inequality since 1700. Princeton University Press, Princeton
- Lindert PH, Williamson JG (2017) Inequality in the very long run: Malthus, Kuznets, and Ohlin 11: 289–295
- Malines van Ginkel HF (1926) Verslag van Den Economischen Toestand Der Inlandsche Bevolking, vols I & II. Landsdrukkerij, Weltevreden
- Meijer Ranneft JW, Huender W (1926) Onderzoek Naar Den Belastingdruk Op de Inlandsche Bevolking. Landsdrukkerij, Weltevreden
- Milanovic B (2016) Global inequality: a new approach for the age of globalization. Harvard University Press, Cambridge
- Milanovic B (2018) Towards an explanation of inequality in premodern societies: the role of colonies, urbanization, and high population density. *Econ Hist Rev* 71(4):1029–1047
- Milanovic B, Lindert PH, Williamson JG (2011) Pre-industrial inequality*. *Econ J* 121(551):255–272
- van Nederveen Meerkerk E (2019) Women, work and colonialism in the Netherlands and Java. Comparisons, contrasts and connections. Palgrave Macmillan, London, pp 1830–1940
- Nugent JB, Robinson JA (2010) Are factor endowments fate? *Revista de Historia Económica Journal of Iberian and Latin American Economic History* 28(1):45–82
- Pelzer KJ (1978) Planter and peasant: colonial policy and the agrarian struggle in East Sumatra 1863–1947. Brill, Leiden
- Piketty T (2014) Capital in the twenty-first century. Harvard University Press, Cambridge
- Reys RJW (1925) De Inkomstenbelasting Der Inlanders En Met Hen Gelijkgestelden in Nederlandsch Oost-Indië. Nijhoff, The Hague
- Ross C (2017) Developing the rainforest: rubber, environment and economy in Southeast Asia. In: Economic development and environmental history in the anthropocene: perspectives from Asia and Africa. Bloomsbury, London, pp 199–218
- Scheidel W (2018) The great leveler: violence and the history of inequality from the stone age to the twenty-first century. Princeton University Press, Princeton
- Scheltema AMPA (1927) De Ontwikkeling van de Agrarische Toestanden in de Priangan. *Landbouw* 3(271–304):317–368
- Sokoloff KL, Engerman SL (2000) History lessons: institutions, factors endowments, and paths of development in the new world. *J Econ Perspect* 14(3):217–232
- Tjhan TS (1933) De Landrente-Belasting. N.V. Drukkerij De Banier, Rotterdam
- Touwen J (2001) Extremes in the archipelago. Trade and economic development in the outer Islands of Indonesia, 1900–1942. KITLV Press, Leiden
- van der Eng P (2004) Productivity and comparative advantage in rice agriculture in South-East Asia since 1870. *Asian Econ J* 18(4):345–370
- van der Eng P (2010) The sources of long-term economic growth in Indonesia, 1880–2008. *Explor Econ Hist* 47(3):294–309
- van Leeuwen B, Földvári P (2017) The development of inequality and poverty in Indonesia 1932–2008. *Bull Indonesian Econ Stud* 52(3):379–402
- von Benda-Beckmann F, von Benda-Beckmann K (2013) Political and legal transformations of an Indonesian polity. The Nagari from colonisation to decentralisation. Cambridge University Press, Cambridge
- Wellenstein EP (1926) Het Rapport Meyer Ranneft—Huender Nopens Den Belastingdruk Op de Inheemsche Bevolking van Java En Madoera. *Koloniale Studien* 10(1):111–151
- Wertheim WF (1956) Indonesian Society in transition. A study of social change. W. van Hoeve, The Hague and Bandung
- Williamson JG (2015) Trade, growth, and distribution in Southeast Asia, 1500–1940. In: Coxhead I (ed) *Routledge handbook of Southeast Asian economics*. Routledge, New York, pp 22–42

Williamson JG (2011) Trade and poverty. When the third world fell behind. MIT Press, Cambridge

van Zanden JL, Marks D (2012) An economic history of Indonesia. Routledge, London

van Zanden JL, Marks D (2020) Javanese rice prices. 2020. <https://www.cgeh.nl/indonesian-economic-history>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.