Coercive Institutions and Female Empowerment: Legacies of a Colonial Experiment on Java

Simon Felgendreher[1], Ola Olsson[2], Michele Valsecchi[3]

Abstract

What are the long-run implications of coercive colonial institutions on female empowerment? In this paper, we study the short- and long-term effects of a colonial experiment - the Dutch Cultivation System on Java (1830-1870) - that forced the native population to switch from subsistence rice production to sugar and coffee cultivation for exports. First, we show that the cultivation of sugar, but not the cultivation of coffee, is associated with stronger historical women empowerment. An IV strategy, exploiting the suitability of districts to sugar production, provides an even larger (albeit imprecisely estimated) effect. Second, on the basis of the existing literature, we hypothesize that this institutional shock might have affected current female empowerment through: i) An increase in female labor demand (driven by an increase in infrastructure and the emergence of a modern sector). ii) An increase in female labor supply (driven by an decrease in their bargaining power within the household). Consistent with the labor demand channel, we show that sugar and coffee cultivation is associated with greater density of roads and railways, which, in turn, are associated with a several measures of female empowerment. However, contrary to the predictions of the labor supply channel, we show that sugar cultivation, while having a negative effect on the historical male to female ratio, does not affect current female empowerment indicators after controlling for historical infrastructure.

Keywords: Coercive institutions, sugar cultivation, women empowerment, sex ratio.

JEL Classification: N25, O1, O12

[1] Gothenburg University, Department of Economics, Vasagatan 1, 405 30 Gothenburg, Sweden, e-mail: simon.felgendreher@economics.gu.se
[2] Gothenburg University, Department of Economics, Vasagatan 1, 405 30 Gothenburg, Sweden, e-mail: ola.olsson@economics.gu.se
[3] New Economic School, 121353 Moscow, 45 Skolkovskoe shosse, Russia, e-mail: mvalsecchi@nes.ru
Acknowledgment: We are grateful for comments from Randi Hjalmarsson, Andreas Madestam, Melissa Rubio, Pablo Selaya and from seminar participants at University of Gothenburg and the ASWEDE conference in Örebro. Thanks to Louise Granath, Josef Kjellander and Carl Nolander for excellent research assistance.

1. Introduction

It is by now well recognized that coercive or extractive colonial institutions during the last five centuries often led to adverse economic outcomes that have persisted to the present day. In early seminal works such as North (1990), Engerman and Sokoloff (2000), Acemoglu et al (2001, 2002), it is analyzed how specific geographical and production-related conditions in colonized countries gave rise to a great diversity in colonial strategies by the colonizing powers and their successor states, which in turn became manifested in very different contemporary institutional configurations in terms of property rights, political rights, and economic inequality.

A less often studied issue concerns the impact of coercive colonial institutions on female empowerment. Several studies have addressed how historical shocks such as the rise of the Neolithic (Hansen et al, 2015), the use of the plough (Alesina et al, 2013), the slave trade in Africa (Teso, 2017), male military enrollment during World War II (Acemoglu et al, 2004) and the introduction of market incentives for tea cultivation in China (Qian, 2008) have affected various dimensions of gender balance in society, but none of these studies have explicitly focused on the legacies of coercive colonial policies or institutions.[1]

In this research, we add to the literature by studying the long-term implications on female empowerment of a particular colonial institution; the Cultivation System practiced on Java in 19th century Dutch East Indies. The key element of the Cultivation System (1830-1870) was the forced cultivation of sugar and coffee for exports to Europe. This "colonial experiment", hastily rolled out in the wake of a deep Dutch government crisis, forced the native population in certain localities to switch from subsistence rice production to either coffee or sugar cultivation under an extensive price-fixing scheme. We argue that the cultivation of sugar and coffee constituted somewhat similar treatments when it comes to the development of infrastructures and the modernization of the economy, while they constituted two different types of "treatment" in relation to direct female empowerment. The rise of the new export industry was also accompanied by large investments in railways and roads that facilitated the commercial production of both crops. However, the cultivation and harvest of sugarcane was very heavy work, mainly carried out by male laborers who had to endure significant hardships that in turn led

[1] One of few exceptions is a recent paper by Grosjean and Khattar (2015) who study the long-term implications of the male biased sex ratio in Australia resulting from the British deportation of male convicts during the colonial era. We will discuss that paper further below.

to a female-biased sex ratio in the sugar districts among the native population. Coffee cultivation was not at all as heavy and women often had a comparative advantage in picking coffee berries. In coffee areas, the sex ratio did not become biased in any direction. None of the crops are important export industries today.

We look at historical measures of female empowerment and find that sugar cultivation is positively associated with the share of divorced females and the share of households led by a female, and negatively associated with the share of married females. The IV estimates (based on the suitability of districts for sugar cultivation) are even larger (although somewhat noisier) and seem to confirm such a relationship. When we analyze whether historical sugar cultivation also has an effect on contemporary outcomes, we do not find any consistent effects on similar measures of empowerment.

We then attempt to separate two potential channels of transmission: a labor demand channel and a labor supply channel. The labor demand channel works as follows: The sugar cultivation might have led to an increase in infrastructures, which might have increased (female) labor demand, and therefore female empowerment. Consistent with this channel, we find that forced sugar cultivation is associated with 1884 roads and railway density and that the latter are positively associated with a wide range of current indicators of female empowerment. Separately from this channel, the labor supply channel works as follows: Sugar cultivation, which decreased the male to female ratio, decreased the household bargaining power of women in the short run by decreasing their likelihood of marriage and forced them to participate in the labor market. In the long run, theory suggests that such an increased activity outside the household might lead to a greater emancipation of women through changes in culture and norms. Contrary to this hypothesis, we find that, while sugar cultivation is associated with a lower historical male to female ratio, it does not lead to greater female empowerment today. Hence, even if the introduction of sugar cultivation implied a shock to local sex ratios, there are no clear signs of a persistent effect on female empowerment today other than through proximity to transport infrastructure.

Our paper is related to a number of recent papers on women empowerment.[2] The most closely associated paper in this tradition is probably Grosjean and Khattar (2015) who study the long-run

[2] Classic references include Becker (1991), Angrist (2002), and Chiappori et al (2002). See Giuliano (2017) for a recent overview of this growing literature. We discuss some of this work further below in the theoretical section.

impact of the large inflow of male convicts to Australia in the 1800s. This colonial shock led to a highly male-biased sex ratio which, according to classical household bargaining models, implied a stronger bargaining advantage for women, which resulted in more leisure and less female labor force participation. The authors show that even today, when the sex ratios in Australia have become balanced, female labor force participation is lower in areas that once had many convicts. Grosjean and Khattar (2015) attribute this effect to cultural norms from the 1800s that have persisted to this day.

Unlike Grosjean and Khattar (2015), our results are not consistent with a cultural model whereby norms from colonial times have lasted until today, but rather suggest that "hard facts on the ground" such as infrastructure is the main legacy from the Cultivation System on modern female empowerment. A plausible explanation is that on the extremely densely populated island of Java with its developed infrastructure, it was harder for cultural biases from colonial days to survive than in more isolated communities in less densely populated countries like the United States or Australia.

The Cultivation System has been previously studied by a number of scholars specialized on Dutch colonial history, including Elson (1984), Boomgard (1989), van Niel (1990), Fasseur (1992), van Zanden (2010), and Bosma (2013). Many of these contributions have focused on the negative short-term effects of the Cultivation System on the Javanese population such as widespread famine, the breakdown of traditional authority, and the coercive nature of production.

In contrast, a recent paper by Dell and Olken (2017), written in parallel with our project, stresses the long-run positive effects of the sugar production of the Cultivation System. Using a hand-written list of sugar producing villages from a government inquiry, Dell and Olken (2017) employ GIS methods to locate the probable location of sugar factories that were the center of sugar processing at the time. They then match the location of colonial sugar factories to the location of villages with contemporaneous data on a number of outcomes such as labor force participation, access to roads, etc. The basic finding in the paper is that former sugar villages are currently better off in most dimensions. Dell and Olken (2017) attribute this development to a proto-industrialization of the Javanese economy as a result of the Cultivation System which turned out to be beneficial in the long run.

Although our paper studies the same colonial institution as Dell and Olken (2017), our paper differentiates from theirs in a number of ways. First, our research question focuses on female

empowerment whereas they primarily analyze the impact on economic development in a broader sense. Second, unlike Dell and Olken (2017), our study compares the effects of sugar with the impact of coffee cultivation. In so doing, we are able to identify not just the legacy of forced sugar cultivation but differentiate between sugar and coffee cultivation, i.e. two crops that required very different agricultural technologies with very different implications for female empowerment. Third, our data is gathered from official annual colonial production statistics on district level (Kolonial Verslaag, 1868, and Algemeen Verslaag, 1871) rather than on hand-written notes from the Umbgrove Commission as in Dell and Olken (2017). Fourth, whereas Dell and Olken's key empirical strategy uses a border RD design comparing sugar and non-sugar villages on a limited part of Java, our identification strategy relies on developing exogenous indicators for the land suitability for sugar and coffee cultivation as instruments for actual coffee and sugar production for all of Java, and then interact these measures with distance to ports, using a strategy similar to Dippel et al (2017). In this sense, we believe our approaches complement each other in a number of ways. Furthermore, our basic finding of a non-persistent direct effect of colonial cultivation on economic outcomes for men and women alike, sets our study apart from Dell and Olken (2017) where the main finding is a beneficial legacy of the Cultivation System on micro level economic development.

Our paper is also related to a growing literature on the long-run effects of colonial infrastructure in developing countries (see for instance Jedwab and Moradi (2016) on colonial railways in Ghana or Donaldson (2010) on colonial railways in India) to the particularly harmful legacies of the cultivation of sugar in former colonies (Dippel et al, 2017), and to research focusing on biased sex ratios and the issue of "missing women" in Asia (Sen, 1990 and Qian, 2008). However, unlike in other parts of Asia, the key bias in our historical setting is "missing men".

Finally, it might be argued that our paper makes a contribution by reporting a non-persistent effect of a major historical shock such as the Cultivation System. Many of the papers in this tradition focus on persistence and we suspect there is a bias in the literature towards results showing some form of persistence. Throughout history, it was most likely the case that several historical shocks that were deemed to be of fundamental importance at the time eventually turned out not to have lasting effects. If research on such episodes are not reported in research articles,

there is a risk of a bias towards "persistence papers" that might obscure our understanding of when history matters and when it does not.

The paper is structured as follows. Section 2 presents the historical background of the Cultivation System whereas section 3 outlines the conceptual framework for the empirical study. Section 4 discusses the data. The main empirical results for the historical and contemporaneous outcomes are presented in section 5. Section 6 concludes.

2. Historical background

In this section, we first give a brief overview of the colonial history of Java, whereupon we elaborate upon the key colonial institution that we hypothesize had a lasting effect on female empowerment in Javanese society; the Cultivation System of 1830-1870.

2.1 Colonial history of Java

The (Dutch) United East India Company, the VOC (Vereenige Oost-Indische Compagnie) was founded in 1602 with its headquarters in Amsterdam. The first permanent post on Java was established in 1611 in Batavia (today Jakarta) which also became the administrative center for the VOC. At first, only the hinterland of the trading post was controlled by the VOC, but its influence grew gradually by conquests and alliances with local rulers. In the mid-18th century, the entire Western part and the Northern coast was conquered by the VOC. In 1800, the VOC went bankrupt due to mismanagement and was dissolved. Its entire territories were transferred to the Dutch government and became its formal colony, the Dutch East Indies.

With the establishment of Herman Willem Daendels as Governor General in 1808 in Batavia, the installation of a formal colonial administration started. In 1811, Java was conquered by the British as a consequence of the Napoleonic wars in Europe, but handed back to the Dutch after a short period in 1816. In 1825, a rebellion known as the Java War started, due to the discontent of the population with the bad living conditions. It spread over large parts of Central and East Java, but was eventually defeated entirely in 1830 by the Dutch. This was the last major uprising by the population before the struggle for independence and the whole island of Java fell under more or less direct control of the Dutch for the first time (Ricklefs, 2008).

2.2 The Cultivation System

General characteristics

With the control over the whole island in 1830, the Dutch were pressed to raise more revenues from their colony. After Belgium's secession from the Netherlands in 1830 and the costly Java War, the Dutch government was close to bankruptcy and desperately needed new sources of income. In this situation, the King appointed his trusted advisor Johannes van den Bosch as Governor-General (1830-34) of the colony, known for his somewhat radical ideas on colonial agricultural policy. In the words of Bosma (2013, p 91-92):

> "The forced Cultivation System was the brainchild of Johannes van den Bosch…He was convinced the Javanese peasants rarely needed a full day's work to make a living and pay land rent. While rejecting slavery as a system, he did not mind using forced labor."

In order to better exploit the Dutch East Indies, the Cultivation System (*Kultuurstelsel*) was quickly introduced in 1830 and stayed in place until 1870. It was based on the "land rent" introduced during the British interregnum where villages should pay 40 percent of the crops produced as a tax to the colonial government. In practice, this tax was hard to implement and real tax payments were often much lower.

To circumvent the problems with the former land rent, under the Cultivation System each village should reserve a share of its land to cultivate crops for the government. The crops would then be sold to the colonial government at a fixed price. With the income produced by these crops, the village would then be able to pay its financial commitments from the land rent. If they earned more than they were required to pay, they would be able to keep the surplus. If the opposite was the case, they would have to raise the revenue from other sources (Ricklefs, 2008).

In practice, the Cultivation System was not a uniform system. The introduction of cultivations began in all residencies in a relatively arbitrary way (Fasseur 1992, p.31).[3] There was no documented master plan from van den Bosch or anyone else prescribing exact principles that colonial administrators should follow when hastily rolling out the Cultivation System after 1830. However, as we will show empirically later in the paper, both sugar and coffee cultivation was largely determined by the suitability of the soil to grow the respective crop and the distance to the

[3] For instance, Bosma (2013, p 98) cites a liberal critic van Soest who already in 1869 claimed that there was no rationale whatever for making Pasuruan in Eastern Java a center for sugar production.

next port. Especially for sugar, which was produced in larger quantities than coffee during the Cultivation System, transport costs on land were very high and therefore it was important to minimize the distance to the next port.

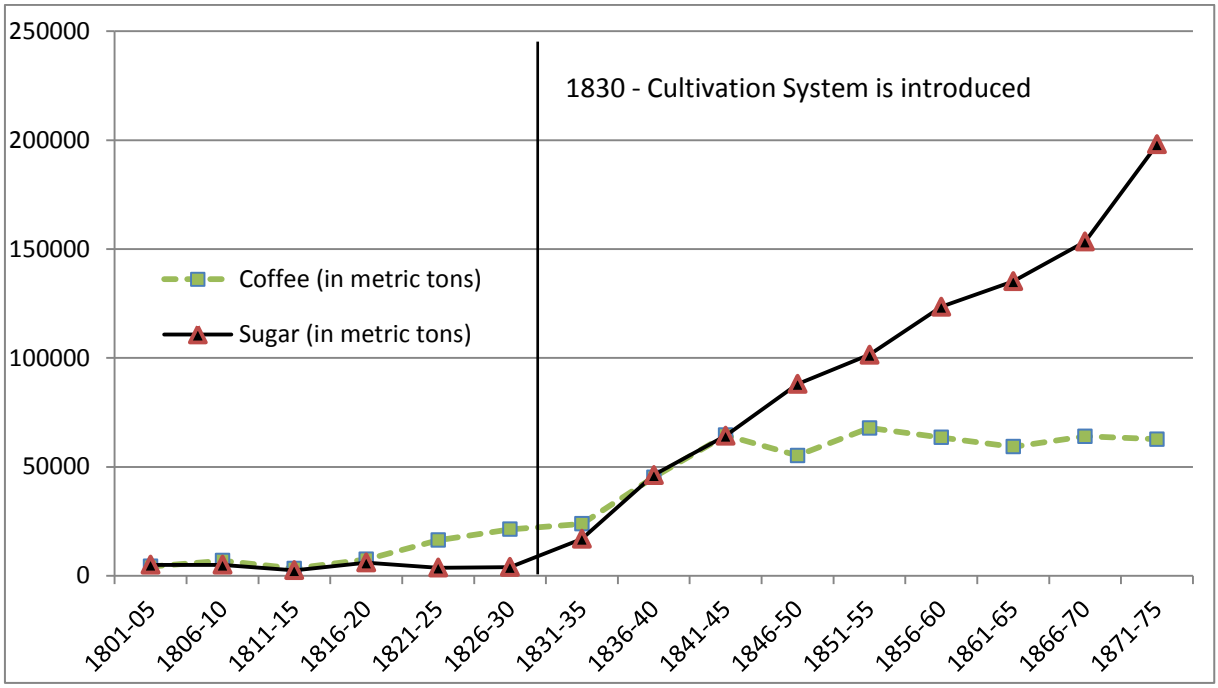
The planning of plantations was done by the director of cultivation in Batavia and the Residents (regional directors) were instructed to follow the orders. The Resident had normally a lot of leeway in executing the orders, since the director of cultivation knew little about the exact conditions in each Residency (Fasseur 1992, p. 45). As a consequence, the intensity of the Cultivation System and its impact on the local population varied widely across Java. A key role in implementing the system was assigned to the native structures of administration (Sutherland 1979, p. 11-13). Especially the village heads were incentivized to ensure the cultivation of the export crops. They received 8 percent of the total amount paid to the village (Fasseur 1992, p.39).

With the introduction of the Cultivation System, the colonizers experimented with various crops using forced cultivations, among others sugar, coffee, indigo, tea and tobacco. For all but coffee and sugar, the production did not prove to be profitable in the long run (Bosma 2013, p.94). In the period 1850-59, revenues from coffee and sugar stood for more than 90 percent of the profits from the Cultivation System. During the same period, revenues from the Cultivation System made up over 30 percent of the total budget of the Netherlands and staying at a similar level until the end of the Cultivation System in 1870 (Fasseur, 1992, p. 150-151). Van den Bosch's forced labor regime had thus in a relatively short time made the Dutch East Indies a fiscal success which had saved the weak Dutch public finances from bankruptcy.

Sugar cultivation

The development of sugar and coffee production on Java during the period 1801-1875 is shown in Figure 1. Before the introduction of the Cultivation System, there already existed some sugar plantations on Western Java. These cultivations were established on privately owned land of Europeans and used wage labor. They were centered south of Batavia, in an area that did not continue to produce sugar under the cultivation system (Bosma, 2013, p. 14, 89). Before 1830, total sugar production amounted to on average about 5000 tons.

Figure 1: The development of sugar and coffee production for export on Java, 1801-1875



Note: Graph is based on data from Boomgaard (1989). The unit on the horizontal axis is metric tons of production.

Immediately after van den Bosch's roll-out of the Cultivation System in 1830, sugar production started to increase very rapidly. Already by 1835, production had grown by 400 percent. In 1875, output was almost 200,000 tons, i.e. fifty times higher than before the forced labor regime. Production continued to increase even after the termination of the Cultivation System in 1870.

Burden on the population

Sugar was planted in rotation with irrigated rice. In order to prepare the fields for sugar cane, they had to be entirely shifted: dikes had to be removed and the land had to be drained. When the fields were returned for rice planting, the whole structure had to be rebuilt (Elson 1990, p. 27-28; Bosma, 2013, p.114-115). After the preparation of the fields, a typical cycle for sugar cane cultivation included planting, weeding, cutting, haulage onto carts, transportation to sugar factory, processing of sugar cane in the factories, and lastly transportation to ports for transport by ships to Europe. The harvesting and haulage of sugar cane exhibited a great burden on the workers and required a lot of physical strength. Once the sugar cane was cut, it deteriorated rapidly and therefore had to be brought to the factories and processed within 48 hours (Bosma,

2013, p.23). Thus, mainly men were employed in the physical work of preparing the fields, harvest and haulage, whereas women and children mainly did the planting and weeding (Bosma, 2013, p. 187). In general, most accounts suggest that more men were employed in sugar cultivation, whereas women and children cultivated crops for subsistence (Elson 1984, p.85).

There are strong indications that the Cultivation System caused deterioration in human welfare among the affected native population. In 1844, a large part of the sugar producing Cirebon area was struck by a severe famine and another one hit Central Java in 1850 (Fasseur, 1993, p 74).[4] During 1849-50, the population size declined for the first time since the war year 1830 (Fasseur, 1993, p 116). Van Zanden (2010, Figure 2) show that consumption per capita on Java decreased substantially from around 1840, particularly during 1842-1855 when levels fell by about 1/4, and consumption levels did not return to pre-reform levels until the 1880s. Van Baten et al (2013) find that average stature among both men and women declined sharply during the second phase of the Cultivation System. Among males born in 1870, average male height in adult age was only 158 cms, a fall in 2-3 cms from the early years of the 19th century. Van Baten et al (2013) attribute this decline in stature to a deteriorating nutritional status of the Javanese, most likely as a result of the forced cultivation of export crops such as sugar and indigo at the expense of local rice production. Average height then increased somewhat after 1870 but remained low throughout the rest of the 1800s. Around 1880, a cattle pest resulted in a loss of about 1/5 of all cattle on the island. Cholera outbreaks in 1864-1875 and 1883-1896 further decimated the local population. Van Baten et al (2013) argue that the integration of Java into global trade probably contributed to this increase in the spread of infectious diseases.

Data on the exact number of workers disaggregated by men and women are not available for the period of the Cultivation System. For the year 1920, around 70 percent of the labor force employed in sugar cultivation was male (Bosma, 2013, p.176). Since there was considerable technological progress in the sugar industry after the end of the Cultivation System, it can be assumed that the physical requirements of workers had been rather higher under the Cultivation System and thus the demand for male workers was also higher compared to 1920. For instance,

[4] Fasseur (1993, p 74) writes: "The famine in Cirebon – and the one in Central Java around 1850 – would probably not have assumed such serious proportions if the population had not already been totally impoverished by the excessive increase in land rent, the market tax, and other taxes. On top of that, they had been exposed to outrageous maladministration and extortion from their own chiefs, against which the European administration had not taken sufficient action. Moreover, several government cultivations had been a heavy burden." A Dutch traveler to Java in 1847 attributed the "poverty, misery and famine" to the "excessive expansion of indigo cultivation".

narrow-gauge trains were gradually introduced to transport the cane from the fields to the factories that alleviated some of the burden related to the transport of the cane.

The intensification of the colonial rule during the Cultivation System and the following years was accompanied by major infrastructure investments. Especially the construction of intercity roads was prioritized to decrease transportation costs of the export crops. Also, starting in 1863, the construction of railways began on the island (van Zanden 2004; Ricklefs 2008; Bosma 2013). An overview of the intercity roads and railways by 1884 is shown in Figure A.2.2. in the appendix. An important fact is further that the colonial roads and railways are still major routes of transportation in modern Java, as can be seen in figure A.2.3.

Coffee cultivation

As shown in Figure 1, coffee production also increased a lot after the introduction of the Cultivation System but the increase was not as dramatic as the increase in sugar output. By the 1840s, coffee production stabilized at a level that was about 200 percent higher than before the new regime.

In general, sugar cultivation made heavier demands on the local population than coffee. Coffee cultivation left rice cultivation relatively undisturbed. Coffee was time-consuming, but it was not as physically demanding (Fasseur 1992, p. 35). The process of harvesting coffee beans has many semblances with picking tea leaves and requires a good eye for ripe beans as well as dexterity and good eye-hand coordination.

Unlike sugar cane, coffee trees were mainly planted in mountainous areas at higher altitudes. The area under cultivation normally did not overlap with the land used for rice cultivation (Clarence-Smith 1994, p.244; Elson 1990, p.27). As a consequence, people were often required to commute to the coffee plantations or live there temporarily. Also the period of preparation of coffee tended to conflict with the rice harvesting season (Elson, p. 27-29). Another form of coffee cultivation was the so called 'hedgerow coffee' that was planted close to the village. According to Clarence-Smith (1994) an advantage of this type of cultivation was that it "could be easily worked by women and children" (p.255). In general, it seems as if women were involved to a large extent in the cultivation of coffee, especially during harvest and transporting the beans back to the village (Clarence-Smith 1994, p.255-257).

End of the Cultivation System

Although the Dutch government earned a lot of revenue from the Cultivation System on Java, it was controversial almost from the start. From about 1848, Dutch liberals in parliament urged the colonial authorities to end the forced labor regimes and proposed an increased use of voluntary wage labor. In 1860, the former colonial official E.D. Dekker published the novel *Max Havelaar* which gave a devastating account of the oppressive and corrupt colonial regime on Java. The Cultivation System was eventually phased out with the introduction of the Agrarian Law of 1870. It led to the gradual abolishment of the monopoly of sugar and coffee growing by the government (Ricklefs 2008; Sutherland, p.13-14).

However, the abolishment of the government monopoly and of forced labor did not end sugar production. The sugar industry remained profitable and there was a strong continuity in production patterns also after 1870. Technological progress in the transport and processing of cane and the introduction of wage labor implied that a proto-industrialization of the Javanese economy took place during the so called "liberal period" 1870-1900. However, although wage labor became the norm during these years, more or less coercive labor supply remained a common problem in villages that continued to be dominated by local elites that often had benefitted a lot during the Cultivation System. Coalitions between village heads and sugar factories prohibited Javanese sugar workers from attaining a stronger bargaining position on the labor market (Bosma, 2013, p 129). Throughout the late 1800s and early 1900s, Javanese cane sugar production amounted to 10-20 percent of world production (Bosma, 2013, p 173).

An increased competition from beet root sugar production, the Japanese occupation during World War II, and the transition to an independent Indonesia eventually made the sugar industry collapse. Despite efforts from Indonesian presidents Sukarno and Suharto, the sugar complex was not successfully revived during independence.

3. Conceptual framework

In this section, we outline the conceptual framework that underpins our empirical investigation. Our basic model is summarized in Figure 2 with time and key historical events on the horizontal axis and our two main hypotheses presented with boxes and causal arrows. The figure has a starting point in 1816 when Java was handed back to the Dutch from the British and the colony Dutch East Indies was formed.

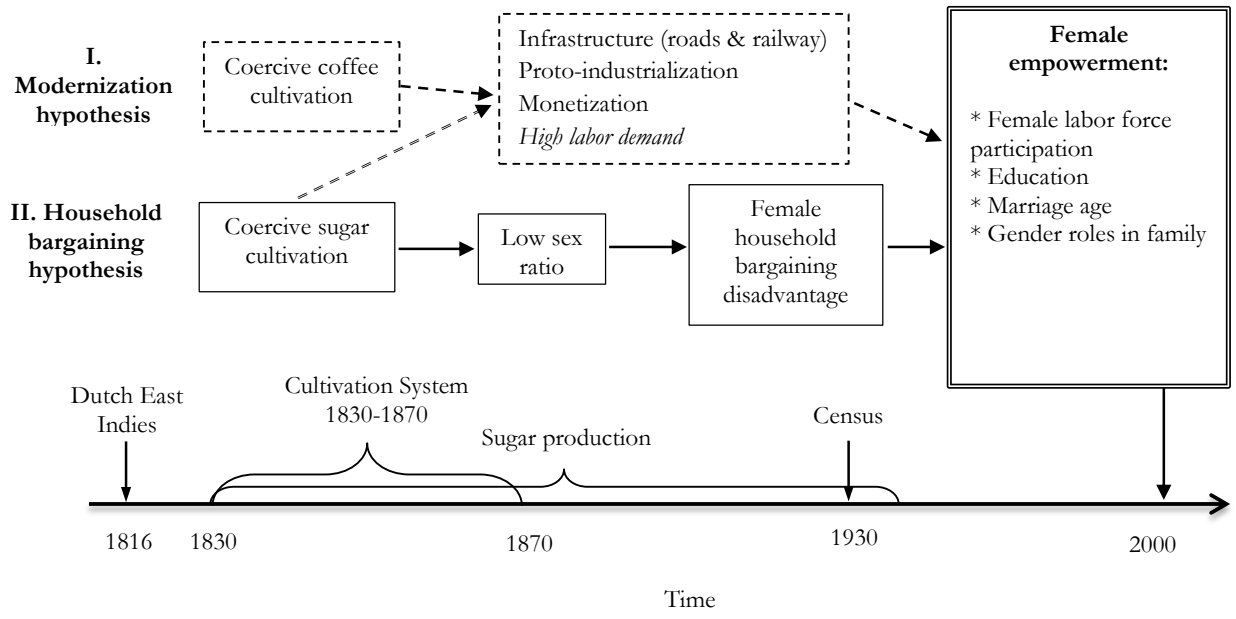
We argue that there are two main causal channels through which the Cultivation System might have affected female empowerment in the short and medium term. The first is the *modernization hypothesis*, proposing that the forced reorientation of the Javanese economy from subsistence rice production to cash crops exports after 1830 eventually led to a proto-industrialization that gave both men and women more access to modern production technologies and job opportunities. More specifically, we consider the possibility that the Cultivation System (i.e. the forced cultivation of both sugar and coffee) led to a monetized system of exchange and to capital investments in infrastructure such as roads and railways which facilitated wage labor outside the home village. Given the "bulkiness" and weight of sugarcane and the speed with which it needed to be brought to factories and ports, we propose that the link between production and railroad investments was stronger for sugar than for coffee. Sacks of coffee from highland areas were most likely more easily transported by road. We analyze effect of this modernization persisted even after sugar and coffee had ceased to be major export crops.

Dell and Olken (2017) emphasize these types of effects and claim that sugar cultivation in the Cultivation System actually was beneficial to economic activity in the long run. In such a modernized economy with a greater general *labor demand*, we might expect that women could more easily find work, regardless of the prevailing sex ratio or the agricultural technology that was in use historically. A greater female labor force participation might in turn be accompanied by a slow change in attitudes whereby the fact that girls observed their mothers taking part in the modern economy gradually led to a greater acceptance of women working and making independent decisions (Fernandez, 2013). In this sense, the modernization of the economy in the late 1800s might have contributed to a greater contemporary female empowerment.

The second causal channel is referred to as the *household bargaining hypothesis* and is illustrated in the middle of Figure 2. This causal explanation focuses on the influence of the Cultivation System on within-household bargaining processes which resulted in an increased female *labor supply*. At the heart of this causal channel lies the new agricultural technologies that the Cultivation System gave rise to. The introduction of forced cultivations of sugar and coffee were both integral parts of the crop export and price fixing regime of the Cultivation System, but we argue that they should still be considered as two different types of interventions due to the very different character of the technologies involved. Sugar was associated with very heavy, mainly male work efforts, substantial capital requirements in order to set up and run the sugar factories,

and a semi-industrial type of organization. Planting and harvesting coffee was not at all as physically demanding as sugar, typically involved women as well as men, and did not require large capital investments in factories.

Figure 2: Two hypotheses concerning the long-run impact of the Cultivation System on female empowerment



Note: The figure shows a timeline with key historical events 1816-2008 on the horizontal axis and the two main hypothesized causal effects in the upper part. Dashed arrows and boxes illustrate the *modernization hypothesis* whereas solid line boxes and arrows represent the *household bargaining hypothesis*. Both are hypothesized to have a persistent effect on indicators for contemporary female empowerment such as female labor force participation, education, and family planning decisions.

In the literature on the long-run impact of sugar plantations in the American colonies, it has been documented by several scholars that sugar plantations were associated with economies of scale and large rent extraction by a planter elite, leading to substantial inequalities between planters and the great majority of African slave workers. These patterns persisted throughout history even after slave emancipation and contributed to the extractive institutions and weak property rights to broad segments of society that can still be observed in several of the affected countries (Engerman and Sokoloff, 2000; Acemoglu and Robinson, 2012; Dippel et al, 2017). In the American colonies, for instance in the Caribbean, the native Indian population typically

perished due to illness and the hardships of plantation work and needed to be replaced by slaves. Sugar cultivation in Asia was mainly carried out by indigenous populations and this was the case also on Java.

The focus in our analysis is the effect of the Cultivation System on a less often studied outcome in the colonial literature; gender imbalances and, more specifically, female empowerment. We hypothesize that the sugar plantation arm of the Cultivation System in Figure 2 had a strong negative effect on the survival rate on Javanese men working in sugar production. As discussed above, sugar production during the Cultivation System and the "Liberal period" was associated with significant hardship, particularly for the men who had to do most of the cutting, haulage, transporting, and manufacturing of the sugarcane. Although the Cultivation System was officially abolished in 1870, sugar production persisted at least throughout the 1930s.

The heavy burden on men in the affected areas led to a female biased sex ratio in the adult population (males/females<1), implying that the total female population outnumbered the total male population in sugar areas. Available data shows that in Java as a whole, the sex ratio was 0.99 in 1814 before the Cultivation System when the island was administered by the British. The ratio then fell to 0.95 in 1930, only to bounce back to more normal levels at 0.985 in 2008. In 1930, the Cultivation System had been abolished for several decades but sugar production was still ongoing. As we will show below, areas with sugar production had a lower sex ratio than non-sugar areas even at this time. We argue that the social environment and production methods in the sugar areas should still have contributed to the biased sex ratio in 1930.

In coffee areas, we have no reason to believe that intense cultivation gave rise to a low sex ratio. As discussed by Qian (2008), women and children appear to have a comparative advantage over men in picking tea leaves in China and Clarence-Smith (1994) argues that the same logic applies to work on coffee trees. Hence, if anything, we would expect that the forced labor regime of the Cultivation System should have implied a more intense exploitation of female labor in coffee areas than in rice or sugar areas and therefore perhaps a higher mortality among women and a higher sex ratio. Since the work was not at all as physically demanding as sugar, there are however reasons to suspect that the effect on women in coffee areas was not as severe as the effect on male mortality in sugar areas. In general, the empirical evidence of forced cultivation on

labor market and educational outcomes has been found to be negative (Nunn 2008; Nugent and Robinson 2010).[5]

A standard argument in the literature since Becker (1991) has been that the level of the sex ratio crucially affects the workings of the marriage market. Individuals from the sex that is in relatively low supply has an advantage in household bargaining over consumption and leisure, i.e. the two most commonly assumed components of an individual's utility function. When men are the scarce sex so that the sex ratio is low, women's bargaining position is relatively weak since men more easily can find other available women to marry. Hence, the outcome of the bargaining under such circumstances should be that women have to work more and consume less than if they were the scarce sex. Furthermore, several women will not find a partner to marry at all and hence need to enter the labor market in adult age in order to earn a living.

Several previous empirical studies have confirmed that the sex ratio hypothesis indeed appears to hold in different settings. Demographic shocks that either decrease or increase the sex ratio often have far-reaching implications for gender balance. For instance, Acemoglu et al (2004) show that World War II mobilization in the United States led to a larger mobilization of female labor. Abramitzky et al. (2011) find that French men living in regions where mortality during World War I was high were more likely to enter marriage and to marry women from higher social classes. Angrist (2002) shows that selective immigration by males into the US in the early 1900s gave rise to a male-biased sex ratio and that second generation female immigrants married more and had more leisure. Chiappori et al. (2002) extend the analysis of the importance of the sex ratio to also include other "distribution factors" which affect labor market outcomes such as marriage or inheritance laws. These results are all conducive to a household bargaining model where the scarce sex has the upper hand, not only in the marriage market, but also in the joint allocation of leisure and work between household members and in the wider pattern of relative empowerment between men and women.

Studying the impact of the forced immigration into Australia by male convicts in the 18th and 19th centuries, Grosjean and Khattar (2015) demonstrate that the very male-biased historical sex ratio had a long-run impact on culture and attitudes in affected regions so that even today, female

[5] However, Bobonis and Morrow (2014) find that forced cultivation of coffee in Puerto Rico in the 19th century led to an increase in the literacy rate in the sugar growing areas of the island. The authors argue that the forced labor scheme suppressed the wages for unskilled workers and thus lead to a higher skill premium, making education more attractive.

labor supply is lower and women are less likely to have high-ranking occupations. In our study of Java, we expect the reverse to hold true: Sugar cultivation was associated with a greater mortality among men and an overall male deficit, resulting in a low sex ratio, a weak household bargaining position for women and consequently a greater participation of women in labor markets.

As shown in Figure 2, we hypothesize that this early household bargaining disadvantage for women implied that affected households typically increased their female labor supply and that this also increased the cultural acceptance of women working. A related process is modeled and simulated on US data by Fernandez (2013). In her model, women experience an intergenerational learning process whereby the long-run payoffs from working are observed by daughters so that attitudes among women gradually move towards a greater cultural acceptance for female labor. In a similar fashion, we argue that the mid-nineteenth century negative shock to the sex ratio and to female bargaining power in sugar areas, constituted a starting point for such a cultural learning process which eventually resulted in greater female labor force participation and a stronger contemporary female empowerment. Hence, we postulate that the initial bargaining disadvantage in a traditional society turned into a stronger role for women in the household in a more modern economy.

Combining the two causal effects of the Cultivation System discussed above, we thus expect a general increase in female (and male) labor demand as a result of the modernization of the Javanese economy and a relative increase in female labor supply due to the biased sex ratio and household bargaining disadvantage of women in sugar cultivating areas. Both effects should contribute to a higher equilibrium level of women employed in the modern sector in former sugar areas and to a change in related indicators of female empowerment such as education.

4. Data

4.1. Historical data

Data on the cultivation of sugar was collected from the colonial reports that were produced annually by the minister of colonies for the Dutch government (Koloniaal Verslag, 1868). For coffee, data is available in a special report on the coffee cultivation on Java during the Cultivation System (Algemeen Verslag Betreffende de Koffijcultuur Op Java, 1871). Both data sets specify the administrative unit, the districts (districten) from the mid-19th century in which the crops

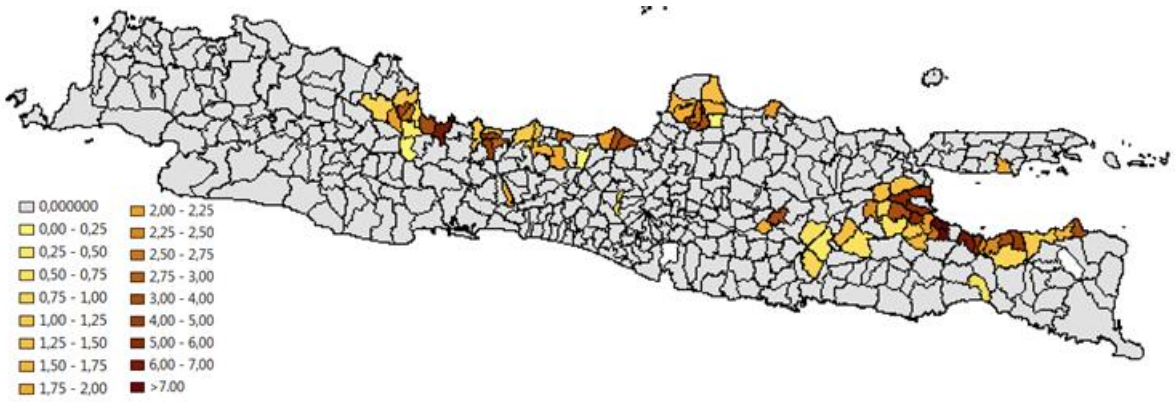
were grown. Using maps from Cribb and Bonne (2010) showing the district boundaries for the year 1857, these districts were identified. The atlas by Cribb and Bonne (2010) also provides maps of the administrative boundaries from 1930. These maps from 1930 were digitized and its administrative units serve as our main unit of analysis in the following. Therefore, the cultivation data and its geographic location from the 1860s are mapped into the districts from 1930.

In 1930, three administrative layers existed on Java (and the neighboring island of Madura). The highest level were the 21 residencies that were divided in 87 regencies and subsequently in 428 districts. These districts are the most disaggregated level on which the data is aggregated and when we refer to "districts" in the following we always mean the districts with the administrative boundaries from 1930.

Population data from 1930 comes from the census carried out by the colonial government (Volkstelling, 1933). Here, population figures disaggregated by sex are available for 425 of the 428 districts. Other variables, such as marriage status, were not collected on such a disaggregated level, but only on the regency level (87 observations). Questions about the religious denomination were not included in the 1930 census. However, Islam had been the dominating religion since the 16th century on Java and most of the rulers had converted to Islam by then. Therefore, it can be assumed that almost the entire native population was Muslim by the time of census (Ricklefs 2008; Michalopoulos et al. 2014).

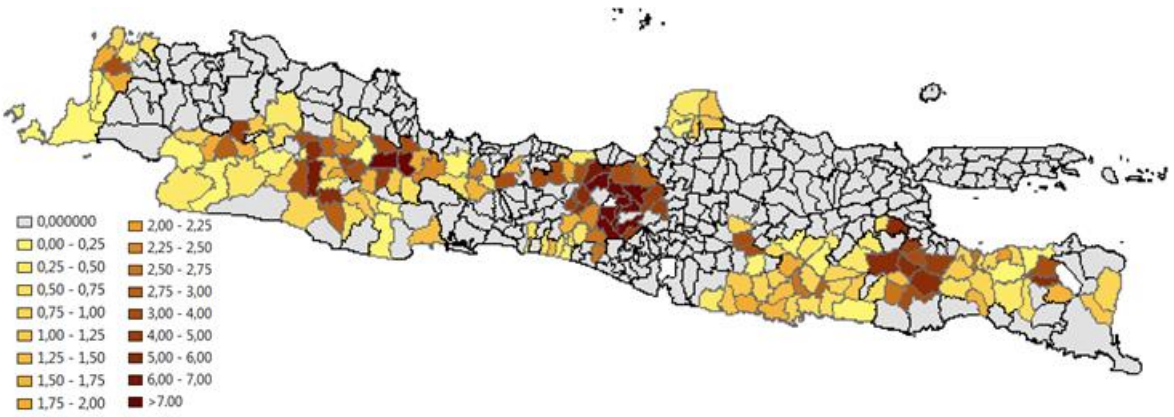
Sugar was mainly grown in the lowlands close to the coast in Eastern Java. Out of the 425 districts, 70 had sugar plantations in 1866 and onwards (see figure 3). Coffee was cultivated in more districts, mainly in the inland and in higher altitudes, (see figure 4). In total, coffee was cultivated in 170 districts. For data on the sex ratio, the first year we have reliable data is 1930 when the census took place. Although this is 60 years after end of cultivation system, there are reasons to believe that differences in the distribution of the sex ratio can be directly linked to crop cultivation during the Cultivation System. The main reason is that forced labor in sugar and coffee cultivating areas de facto continued after 1870 until the end of the 19th century. In Figure 5, the sex ratio in each district is displayed. There is substantial variation in the sex ratio across the island, with many districts having a sex ratio below the “natural” sex ratio of 0.99. Summary statistics of all variables used in this first part of the analysis are shown in table 1.

Figure 3: Map of sugar cultivation 1866



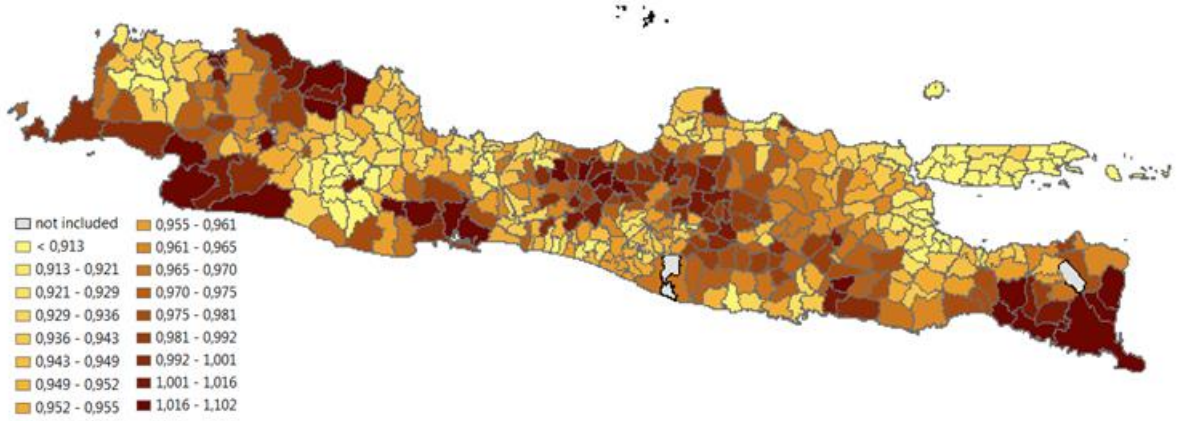
Map shows hectares under sugar cultivation in relation to the total size of the district in square kilometers.

Figure 4: Map of coffee cultivation 1863



Map shows hectares under coffee cultivation in relation to the total size of the district in square kilometers.

Figure 5: Map of sex ratio 1930



Map shows the sex ratio (number of men/ number of women) in 1930 per historical district.

In addition, we use data on railways and roads built during the colonial rule. We digitize a map from 1884 (KITLV, 1884) and calculate the minimal distance to the closest railway and road from the geographical center of each district. Summary statistics of these variables are shown in Table 1 and a map of the railways and roads in 1884 is depicted in Figure A.2.2. in the appendix.

4.2 Geographical data

For our empirical analysis, we assemble a broad range of geographical variables from different sources. Data on elevation and slope are taken from the SRTM30 data set with a resolution of 30 arc-seconds (squares with a length of approximately 900 meters at the equator) and was downloaded from the webpage of webgis. Data on constraints of the soil for crop cultivation comes from the FAO. We use an index that distinguishes between 26 major soil groups and takes the following factors into account to assess the degree of constraints of the soil for agriculture: depth, soil chemical status and natural fertility, drainage and texture. It ranges from one to seven, with a higher number indicating larger constraints of the soil (van Velthuizen et al, 2007). For all these geographical variables (elevation, slope and soil constraints), the mean value for each historical district was calculated and this variable is used in the regression analysis.

In order to determine is suitability of the soil to cultivate sugar and coffee, we followed the approach taken in Dippel et al (2017). For both crops, in a first step a suitability index was constructed based on the temperature, rainfall, soil Ph, elevation and slope of the terrain. The index ranges from 1 to 4 with a higher number standing for a decreasing suitability. In a second

step, the percentage of the area in each district that was suitable for the cultivation of the respective crop (suitability index not higher than 2) was calculated and used as our suitability measure. For more details on the data sources, see appendix A.2.12.

For the construction of the instrumental variables, also the distance to the closest port in Central or Eastern Java was calculated. The main ports in the 19th century were identified using Touwen (2001). Using ArcGIS, the distance to the closest port from the geographical center of each district was calculated. In the same way we calculated the shortest distance to the coastline.

4.3 Contemporaneous data

We use two different survey-based data sets. To control for the underlying socio-economic structure in each village, we use variables from the Village Potential Study (PODES) 2008. For the survey, village heads or members of the bureaucracy were surveyed and it covers all villages and urban centers in Indonesia. For data on the individual level, we use the National Socio-Economic Survey (SUSENAS) 2009. It includes among others data on expenditure, working and health status. For Java and Madura, more than 200,000 people were surveyed belonging to more than 80,000 different households. An overview of the variables used for our analysis from both PODES and SUSENAS can be found in Table 7.

5. Empirical analysis

5.1 Empirical strategy and identification

Our basic empirical strategy will be to exploit the sudden introduction of the Cultivation System by van den Bosch in 1830 as a quasi-random social experiment with two different types of treatment - sugar and coffee cultivation – and a control group of Javanese districts that were not affected. As argued above, the very rapid rollout of the program without any clear guiding principles from the colonial government and the limited knowledge among colonial officials of the local agricultural potential for sugar and coffee production, implied that the process was characterized by a great degree of randomness.

However, in determining the locations where to produce the two main types of crops, it is plausible that colonial authorities did have to take into account basic geographical characteristics, including elevation, distance to ports, and the character of the soil, as well as social and political

circumstances that are unobservable today. Given the presence of such unobserved decision processes, simple OLS results might be biased due to variation in relevant omitted variables that might be present even after controlling for variables such as crop suitability and the distance to the closest port.

Therefore, we will complement our standard OLS analysis with empirical results based on an instrumental variable strategy. Since our main interest is on the effect of sugar cultivation on female empowerment, we will instrument sugar cultivation in the main specifications while still estimating the effect of coffee cultivation with OLS. However, the results do not change for coffee cultivation when it is instrumented, as we will discuss later and show in the appendix. Our key determinants of sugar cultivation are soil suitability and distance to the closest port. While these variables can potentially serve as instruments for sugar cultivation on their own, we use the interaction terms of sugar suitability and distance to the closest port as an instrument in order to address concerns about the validity of the instrument. Next to the interaction term, we also include the individual components as control variables in the regression, namely sugar suitability and distance to the closest port. Thus, the identifying assumption is that the interaction term of sugar suitability and the distance to the closest port does not influence our outcome variables other than via the forced cultivation of sugar during the cultivation system. Since we include sugar suitability and distance to the port as control variables next to our instruments, we are confident that the exogeneity assumption is fulfilled.

By using the interaction terms with the distance to the closest port as one component, the IV strategy is similar to the approach taken in Karadja and Prawitz (2016), where the distance to the closest port is interacted with extreme weather events to predict emigration from Sweden in the 19th century. Also, since all ports were already established before the start of the Cultivation System, we can exclude endogeneity in the selection of these sites (Cribb and Bonne 2010). Thus, the first-stage equation we are estimating is the following:

$$\bar{S}_d = \delta_0 + \delta_1 DistPort_d + \delta_2 SugarSuit_d + \delta_3 SugarSuit_d * DistPort_d + C_d + \upsilon_d \qquad (1)$$

In this equation, the endogenous variable $\bar{S}_d$ is the share of total land area devoted to sugar cultivation in district d in 1866, $DistPort_d$ is the distance from the geographical center of d to the closest port, and $SugarSuit_d$ is the share of land suitable for sugar cultivation in district d. The excluded instrumental variable in the second stage is the interaction term

$SugarSuit_d$*$DistPort_d$ with coefficient δ_3. In addition, we include a vector of control variables C_d, including the population density in 1930, years of colonization, general soil constraints, land suitability for coffee, and average elevation and slope in district d. v_d is the error term.

The descriptive statistics of the main variables observed on district level are shown in Table 1. The key explanatory variables *Sugar* and *Coffee Cultivation* are measured as total hectares under actual cultivation per square kilometer of total district land area in 1866 and 1863 respectively.[6] *Sugar* and *Coffee suitability* are defined in the data section and in the Appendix. *Sex ratio in 1930* is simply defined as total number of males divided by total number of females in the population in 1930 where 1 indicates a completely balanced population.

Table 1: Descriptive statistics of main variables

Variables	**Obs.**	**Mean**	**St. dev.**	**Min**	**Max**
Data set A: Cultivation System and its aftermath					
Area Covered: Entire Java					
Sugar cultivation 1866 (percent of land area)	425	0.42	1.31	0	13.63
Coffee cultivation 1863 (percent of land area)	425	0.98	2.26	0	20.11
Distance to railway 1884 (km)	425	44.67	38.79	0.23	153.08
Distance to road 1884 (km)	425	10.32	10.85	0.04	118.95
Sex ratio 1930 (men/women)	425	0.96	0.04	0.83	1.10
Population density 1930	425	0.47	0.53	0.02	6.49
Years of colonization in 1930 (in 100 yrs)	425	1.58	0.79	0	2.78
Geographical controls					
Sugar suitability (share of land area)	425	0.55	0.42	0	1
Coffee suitability (share of land area)	425	0.22	0.30	0	1
Distance to port (km)	425	88.21	60.05	3.49	324.15
Elevation (meters)	425	299.50	323.25	0	1447.40
Slope (percent)	425	0.12	0.10	0	0.44
Soil constraints index	425	5.00	1.39	0	7

Notes: Each observation is a historical district. The geographical controls for sugar and coffee suitability indicate the share of each district that is highly suitable or suitable for the cultivation of the respective crop, according to the categories shown in the Appendix.

The results from estimating equation (1) are shown in Table 2 below. We start by only including sugar suitability, distance to port, and the interaction term of sugar suitability and distance to port without controlling for additional variables in column 1 and add additional

[6] Since 1 sqkm consists of 100 hectares, the measure is equivalent to percentage of total land area covered by coffee and sugar cultivation.

variables in column 2. In both specifications, the coefficient is highly significant and the related F-statistic is in both cases above to the critical value 10.

Table 2: Instrument relevance in first stage

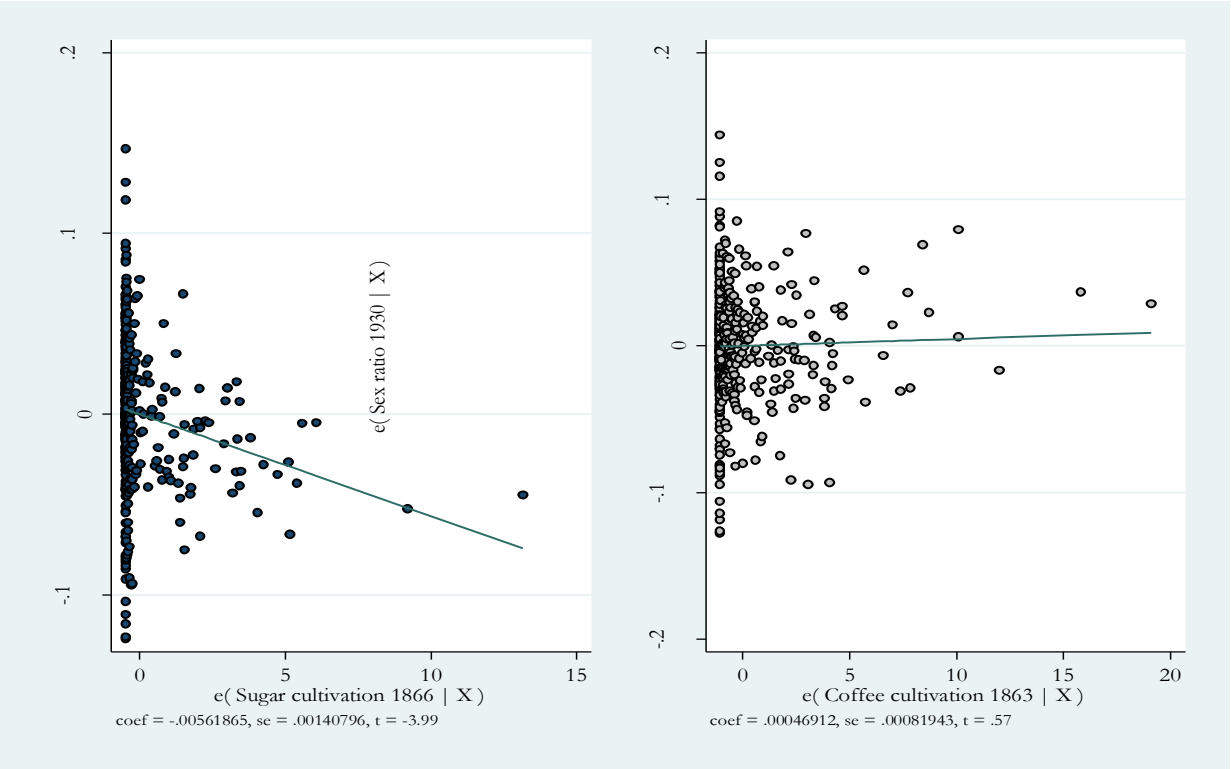
Dependent variable	Sugar cultivation 1866 ($\bar{S}_d$)	
	OLS	OLS
Distance to port	0.001	0.000
	(0.001)	(0.001)
Sugar suitability	1.596***	1.486***
	(0.298)	(0.290)
Distance to port*	-0.010***	-0.009***
Sugar suitability	(0.002)	(0.002)
Mean dep. variable	0.42	0.42
Add. geo. controls	No	Yes
Historical controls	No	Yes
Observations	425	425
F Statistic of instrument	21.90	18.99
R-squared	0.157	0.186

Notes: Each observation is a historical district. Additional geographical controls include coffee suitability, distance to the coast, elevation, slope and general soil constraints. Historical controls include population density in 1930 and the years of colonial rule in 1930. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

5.2 Colonial cultivation, sex ratios and infrastructure

Turning now to the relationship between forced crop cultivation and the sex ratio in 1930, we start by showing the conditional relationships in Figure 6 where we regress the sex ratio in 1930 on the intensities of colonial sugar and coffee cultivation in the 1860s. Sugar cultivation has a pronounced negative association with the sex ratio (fewer men in sugar areas) whereas coffee displays a weak positive slope. The negative relationship with sugar cultivation is significant at the .01-level whereas the estimate for coffee is insignificant.

Figure 6: Conditional relationships between sex ratio in 1930 and colonial sugar/coffee cultivation



Notes: The figure shows the conditional relationships from a regression with *Sex ratio 1930* as the dependent variable and *Sugar* and *Coffee cultivation* as independent variables. Information about estimated coefficients, standard errors and t-values are displayed below the figures.

The full regression results are shown in Table 3. As hypothesized, sugar cultivation is significantly negatively correlated with a lower sex ratio in 1930 across all model specifications, whereas the same is not true for coffee cultivation. At -.003, the point estimates in columns 3-4 seem small. An increase of the sugar cultivation by one standard deviation (1.31) leads to a decrease of the sex ratio by around .004 units. In columns 5-8, the instrument for sugar cultivation is used. The coefficient for sugar increases in absolute magnitude by around six times compared to the OLS estimates. A one standard deviation increase in sugar cultivation is now associated with a decrease in the sex ratio of on average .022 units (or 2.2 fewer men per 100 women). At the same time, the standard errors increase drastically, so that the estimates are only significant at a 5-percent level in columns 5 to 7 and at 10-percent level in column 8. We argue that the results lend support to our hypothesis that the hardships for the male population engaged

in sugar cultivation gave rise to a lower sex ratio whereas no effect was visible in coffee districts. Also, when coffee cultivation is instrumented instead of sugar, there is no negative relationship between coffee cultivation and the sex ratio in 1930 (table A.2.2. in the appendix).

In Table A.1.2, we restrict the sample to areas covered by the 1814 British census, and we test whether sugar-producing and sugar-non-producing areas were similar before the Dutch cultivation system took place. Reassuringly, the estimates suggest that producing and non-producing areas were similar.

Table 3: Sex ratio 1930 and crop cultivation during the Cultivation System

Dependent Variable	Sex ratio 1930							
	OLS				IV Sugar			
Sugar cultivation 1866	-0.004***	-0.004***	-0.003**	-0.003**	-0.019**	-0.018**	-0.018**	-0.017*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)	(0.007)	(0.009)	(0.009)
Coffee cultivation 1863		0.001		0.001		0.001		0.001
		(0.001)		(0.001)		(0.001)		(0.001)
Constant	0.978***	0.978***	0.879***	0.880***	0.987***	0.986***	0.957***	0.953***
	(0.007)	(0.007)	(0.042)	(0.042)	(0.007)	(0.007)	(0.061)	(0.061)
Mean dep. variable	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Mean sugar cultivation	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Mean coffee cultivation	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Distance to port	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Suitability variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo. controls	No	No	Yes	Yes	No	No	Yes	Yes
Historical controls	No	No	Yes	Yes	No	No	Yes	Yes
Observations	425	425	425	425	425	425	425	425
R-squared	0.081	0.082	0.138	0.140				

Notes: Each observation is a historical district. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. Additional geographical controls include distance to the coast, elevation, slope and general soil constraints. Historical controls include population density in 1930 and the years of colonial rule. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4 shows the association between colonial infrastructural investments as observed in 1884 and the intensity of sugar and coffee cultivation in the 1860s. The results in the table should be interpreted with caution since we cannot rule out that more primitive roads were not already in place in the same paths at the time when the Cultivation System was implemented. In any case, the pattern that emerges from the table is that the presence of coffee cultivation strongly predicts close proximity to roads whereas sugar cultivation is strongly associated with proximity to railways. The result also holds when the cultivation of sugar is instrumented for (column 3 and 6). This is consistent with an interpretation that heavy sugarcane loads were suitable for train transportation whereas lighter sacks of coffee beans could be more easily transported by road.

Table 4: Relationship between infrastructure and sugar and coffee cultivation

Dependent Variable	Distance to road 1884			Distance to railway 1884		
	OLS	OLS	IV	OLS	OLS	IV
Sugar cultivation 1866	-0.381*	-0.410*	2.470	-3.066**	-5.117***	-18.644**
	(0.201)	(0.215)	(2.193)	(1.343)	(1.329)	(8.406)
Coffee cultivation 1863	-0.669***	-0.728***	-0.737***	-0.622	-0.897	-0.855
	(0.200)	(0.190)	(0.189)	(0.568)	(0.608)	(0.614)
Mean dep. variable	10.32	10.32	10.32	44.67	44.67	44.67
Mean sugar cultivation	0.42	0.42	0.42	0.42	0.42	0.42
Mean coffee cultivation	0.98	0.98	0.98	0.98	0.98	0.98
Distance to port	Yes	Yes	Yes	Yes	Yes	Yes
Suitability variables	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo. controls	No	Yes	Yes	No	Yes	Yes
Historical controls	No	Yes	Yes	No	Yes	Yes
Observations	425	425	425	425	425	425
R-squared	0.140	0.235		0.092	0.312	

Notes: Each observation is a historical district. The table reports the OLS as well as the 2nd stage estimates of an IV regression with sugar as endogenous variable. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included. Robust standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Furthermore, it is clear from visual inspection that the routes taken by the colonial roads and railways have remained to the modern day. Most of the colonial roads are now modern highways and also railways follow the trajectories of the colonial days. Thus, our interpretation is that colonial infrastructure had a strong persistent impact on modern infrastructure.

5.3 Impact of cultivation system on colonial female empowerment

In this section, we analyze the short-term effects of forced sugar cultivation on outcomes related to female empowerment. Outcome variables come from the 1930 census on Java. Unfortunately, the variables of interest here were only collected on the regency level, the administrative level above the districts under the colonial administration. In total, there were 87 regencies in 1930. Summary statistics are shown in the appendix (Table A.2.1.).

The variables we are interested in are the share of divorced women, households with a female family head and share of married women. When we first look at the correlation between these variables and the sex ratio 1930 in table 5, we see that the coefficients are in line with the predictions from household bargaining models. With relatively few men in a jurisdiction, there are more divorced women, more women are responsible for the household and there are fewer married women. The coefficient is not statistically significant for divorced females. For the other two variables, the coefficients are highly significant, both statistically and economically. A one standard deviation decrease in the sex ratio is associated with an increase in female head of household by around 10 percent and a decrease of married females by around 2 percent.

Table 5: Sex ratio 1930 and female outcomes1930, regency level

Dependent Variable	Divorced Females		HH head female		Married females	
	OLS	OLS	OLS	OLS	OLS	OLS
Sex ratio 1930	-0.055	-0.064	-0.467***	-0.490***	0.242***	0.235***
	(0.061)	(0.077)	(0.069)	(0.091)	(0.064)	(0.081)
Mean dep. variable	0.05	0.05	0.12	0.12	0.43	0.43
Mean sex ratio 1930	0.95	0.95	0.95	0.95	0.95	0.95
Distance to port	No	Yes	No	Yes	No	Yes
Suitability variables	No	Yes	No	Yes	No	Yes
Add. geo. controls	No	Yes	No	Yes	No	Yes
Historical controls	No	Yes	No	Yes	No	Yes
Observations	87	87	87	87	87	87
R-squared	0.013	0.157	0.270	0.403	0.086	0.239

Notes: Level of aggregation is 1930 regencies. Study area is Java and Madura. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10

Table 6 shows the reduced-form effect of forced crop cultivation on the female outcomes along with the infrastructure variables. When OLS is used, the coefficients for sugar cultivation are statistically significant for the share of divorced females and the share of female head of households. Also the sign is as predicted by the theoretical framework. For coffee cultivation, the coefficients are not significant (except for column 6 where it is significant at a 10%-level). When sugar cultivation is instrumented, as shown in the last column for each outcome variable, the estimates increase in absolute terms, but at the same time the standard errors increase drastically, so that the results are no longer statistically significant for sugar cultivation. The distance to a railway or road is not statistically significant in any of the specifications and the point estimates are extremely close to zero. Thus, it seems that any short-run effect on female empowerment of the Cultivation System is associated with the forced cultivation of sugar and in line with the labor supply channel that encourages women to work.

Table 6: Forced crop cultivation and female outcomes 1930, regency level

Dependent Variable	Divorced Females			HH head female			Married females		
	OLS	OLS	IV	OLS	OLS	IV	OLS	OLS	IV
Sugar cultivation 1866	0.012**	0.020**	0.046	0.025**	0.029**	0.204	-0.021**	-0.013	-0.072
	(0.008)	(0.009)	(0.051)	(0.011)	(0.012)	(0.130)	(0.001)	(0.010)	(0.077)
Coffee cultivation 1863	-0.001	-0.001	-0.001	-0.000	0.003	0.005*	0.000	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.000)	(0.002)	(0.002)
Distance to railway 1884		-0.000	-0.000		0.000	0.000*		0.000	-0.000
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
Distance to road 1884		0.000	0.000		-0.000	0.000		0.001	0.000
		(0.000)	(0.000)		(0.000)	(0.001)		(0.001)	(0.001)
Mean dep. variable	0.05	0.05	0.05	0.12	0.12	0.12	0.43	0.43	0.43
Mean sugar cultivation	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Mean coffee cultivation	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Distance to port	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Suitability variables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Add. geo. controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Historical controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Observations	87	87	87	87	87	87	87	87	87
R-squared	0.068	0.250	-	0.026	0.253	-	0.028	0.215	-

Notes: Level of aggregation: 1930 regencies. In the regressions using an instrumental variable strategy, the excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.10

5.4 Contemporaneous outcomes

In order to analyze the long-run impact of forced crop cultivation on outcomes related to female empowerment, such as labor force participation and education level, we estimated regressions based on the following equation:

$$y_{idp} = \alpha + \beta_1 \hat{S}_{dp} + \beta_2 CoffeeCult_{dp} + Z_{dp} + X_{dp} + X^C_{idp} + \delta_p + \varepsilon_{idp} \quad (2)$$

In this equation, y_{idp} is the respective outcome for individual i in historical district d in contemporary province p. $\hat{S}_{dp}$is either the actual level of sugar cultivation (S_{dp}) or the predicted value from the IV-specification ($\bar{S}_{dp}$). $CoffeeCult_{dp}$ is the actual level of coffee cultivation. Z_{dp} are the historical infrastructure variables distance to road and distance to a railway. X_{dp} is a vector of geographical, historical, or contemporaneous controls in district d of province p. In addition, we include contemporaneous individual level controls X^C_{idp} and modern province fixed effects δ_p.

The reason for sometimes including the infrastructure variables in Z_{dp} is that we want to explore what happens to β_1 and β_2 when these other colonial variables from our theoretical framework are included. We know from our analysis above that the distance to historical railways and roads is correlated with sugar and coffee cultivation. If, for instance, β_1 has a significant coefficient when Z_{dp} is not included, whereas β_1 turns insignificant when Z_{dp} is included, this could indicate that all the causal effect of sugar cultivation on the outcome of interest runs through the intermediate channels in Z_{dp}, i.e. via the increased labor demand in areas that have had better access to infrastructure historically. If β_1 remains significant after including Z_{dp}, this suggests that sugar cultivation affects female empowerment via the labor supply channel, where women are more self-reliant and more likely to work.

Standard errors are clustered on the historical district d. Summary statistics for all variables are shown in Table 7 below. Our sample now includes more than 200,000 individuals on Java, about half of whom are women. The first-stage IV-regression is estimated based on equation (3):

$$\bar{S}_{dp} = \delta_0 + \delta_1 DistPort_d + \delta_2 SugarSuit_d + \delta_3 SugarSuit_d * DistPort_d + Z_{dp} + X_{dp} + X^C_{idp} + \delta_p + \upsilon_{idp} \quad (3)$$

As before, the excludable instrumental variable is the interaction term between $SugarSuit_d$ and $DistPort_d$. In contrast to the historical analysis in section 5.2., here the units of observation are individuals rather than aggregated data points on the historical district.

Results when coffee cultivation is instrumented instead of sugar are shown in tables A.2.6 to A.2.9. First stage results are shown in table A.2.3. in the appendix and indicate that both instruments are relevant when used as instruments individually.

We start by estimating the effect on forced crop cultivation on the extensive margin of labor force participation, i.e. whether a person is working or not. The results are shown in Table 8, columns 1-4. The outcome variable in this case is a binary variable. Like in all three tables ahead, the first and second columns include the whole sample and uses OLS whereas the third and fourth columns uses the IV (2SLS) estimator with a sample of only women in the fourth column. In the first column, we estimate the regressions without the other key colonial variables measuring infrastructure and the sex ratio, whereas we include these variables in the second to fourth columns. All specifications below are estimated using all instrument controls, additional controls and province fixed effects. Standard errors are throughout clustered at historical district level. Due to space constraints, we do not report the IV- regressions that are equivalent to the column 1 specification (excluding distance to railway and roads) in the main section, but in table A.2.5 in the appendix. However, these results are basically identical to the results in the third column and therefore not discussed explicitly.

Table 7: Summary statistics of main variables

Variables	Obs.	Mean	St. Dev.	Min	Max	Sample Restriction
Data set B: Contemporaneous variables						
Outcome variables						
Labor market participation (binary)	203,185	0.72	0.45	0	1	Age >18 and <70
Years of education	202,541	7.84	4.06	0	17	Age >18 and <70
Household head female (binary)	87,632	0.15	0.34	0	1	Household heads only
Age at first marriage	92,985	19.16	4.12	10	60	Married females only
Sector Employed (all binary)						
Agriculture	137,031	0.32	0.46	0	1	
Manufacturing	137,031	0.15	0.35	0	1	Working individuals
Services	137,031	0.46	0.50	0	1	only
Other	137,031	0.08	0.27	0	1	
Infrastructure variables						
Distance to railway 1884 (km)	425	44.67	38.79	0.23	153.08	
Distance to road 1884 (km)	425	10.32	10.85	0.04	118.95	
Geographical Controls						
Land suitable for sugar (share)	425	0.55	0.42	0	1	
Land suitable for coffee (share)	425	0.22	0.30	0	1	
Distance to closest port (km)	425	88.21	60.05	3.49	324.15	
Distance to coast (km)	425	29.45	21.11	0.42	84.27	
Elevation (meters)	425	299.50	323.25	0	1447.40	
Slope (in percent)	425	0.12	0.10	0	0.44	
Soil constraints	425	5.00	1.39	0	7	
Historical controls						
Population density 1930 (inhab. per sq. km)	425	466.57	533.00	20.10	6490.86	
Years of colonial rule (in 1930)	425	158.00	78.27	0	278	
Contemporaneous controls						
Population density 2008	425	1439.11	2448.30	125.22	24564	
Sex ratio 2008	425	0.99	0.03	0.89	1.11	
Individual variables						
Female	215,905	0.51	0.50	0	1	
Age	215,905	42.13	15.61	19	98	

Notes: The geographical controls for sugar and coffee suitable land indicate the share of each district that is highly suitable or suitable for the cultivation of the respective crop, according to the categories shown in the appendix. The variable "Labor market participation" is a dummy variable whether an individual is working or not. The variables in the category "Sectoral employment" are all binary.

As in the tables below, the results for the historical levels of sugar and coffee cultivation are sometimes consistent with our hypothesis and sometimes not, and the estimates are not consistently measured with precision. For instance, when *Labor force participation* is the dependent variable in Table 8, the estimates are negative and statistically significant in the OLS columns for *Sugar cultivation* but at a very low economic significance. In the IV, estimates are positive but insignificant. Similar non-robust results are found for sugar and coffee cultivation in the other tables. Our interpretation is that we cannot reject the null hypothesis of no impact at all. In other words, our hypothesis that historical cultivation might still affect female empowerment is not supported by the data. We will return to discuss potential reasons below. Note also that for *Labor force participation*, the colonial infrastructure variables are not significant either, except for the distance to road, where the size of the coefficient is however extremely small.

Regarding education in columns 5-8, we note that average years of education is 7.8 years for men and women together and 7.3 years for women. Focusing on the results in columns 7-8 where the cultivation of sugar is instrumented for, we find that sugar cultivation is not statistically significant. For forced coffee cultivation, the coefficient is positive and significant at a 10%-level. However, when the instrument is used for coffee cultivation, the coefficient is no longer significant (see table A.2.6. in the appendix). Results in column 7 indicate that distance to a colonial railway has a negative association with years of education. When distance to a colonial railway decreases by one standard deviation, years of education increase by approximately 8 months (0.62 years or around 16% of the standard deviation of years of education). The relationship between colonial roads and education is similar but different in magnitude. If the distance to the closest road decreases by one standard deviation, years of education increase by around 6 months (0.5 years or around 12% of the standard deviation of education). The coefficients for women reported in column 8 are basically identical and thus the effect does not differ between men and women.

In Table 9, we investigate next whether there is an effect of our colonial variables on the underlying economic structure. We use data on sectoral employment and classify individuals as either working in agriculture, manufacturing, services or remaining sectors (for example mining) as our outcome variables. Our sample now includes men and women who are taking part in the labor force (137,031 individuals in total including 53,404 women). Neither colonial coffee nor sugar cultivation has any consistent direct impact on whether men or women work in agriculture or in manufacturing. However, what does seem to matter once again is access to railroads and roads. Also men but even more so women are prone to work in agriculture when

distances to colonial roads are large. A similar but opposite pattern emerges for employment in manufacturing in columns 5-8 where *Distance to railway* is negative and significant at the 1 percent level in columns 6-8. The coefficients are small in size but it can be argued that they are still economically significant. For instance, a decrease in the distance to a colonial road by 10 kilometers (one standard deviation) decreases the likelihood to work in the agricultural sector by 10% and increases the likelihood to work in manufacturing by 4% for women. Also work in the service sector increases when the distance to colonial roads decrease, for both men and women, as shown in table A.2.4. These results seem to support our hypothesis of a persistent labor demand effect of colonial infrastructure and that the Cultivation System in this sense appears to have supported the industrialization of Java.

Lastly, in Table 10, we use two indicators for female empowerment as dependent variables: Whether a woman is household head in columns 1-4 and the marriage age of women in columns 5-8. The household head indicator was used also for the 1930 sample in Table 6. At that time, 12 percent of women were household heads whereas the contemporary percentage is 14. We interpret women's marriage age as a proxy for women's degree of independence where a higher age indicates higher independence. The average marriage age in our sample of women only is 19.35 years.

Once again, historical coffee and sugar cultivation typically are weak or inconsistent predictors of female empowerment. A more striking result is the importance of colonial infrastructure for female empowerment. The shorter the distance to a road, the greater the probability that a woman is the head of household. Regarding marriage age, the effect is economically important. A one standard deviation decrease in distance to a colonial railway (38.8 km) is associated with an increase in marriage age of around 0.5 years (around 10% of the standard deviation of marriage age). A one standard deviation in decrease in distance to a road increases the marriage age by around 0.4 years.

Our general interpretation of these results for the contemporaneous outcomes is that there is no consistent evidence in favor of a persistent cultural impact of the Cultivation System on people's modern choices about labor force participation, education, etc. The historical shocks to the sex ratio and to indicators of female empowerment in sugar areas that were evident in 1930 are no longer reflected in a range of outcome variables. However, access to railways and roads that were there already in colonial times and which coincide with modern infrastructure, have a strong impact on contemporaneous economic choices. This last result is consistent with our *modernization hypothesis* of a greater degree of industrialization and a higher labor demand close to infrastructure.

Table 8: Labor market participation and education level

Dependent Variable	Labor Force Participation: (Yes=1/No=0)				Years of education			
	OLS Men and Women	OLS Men and Women	IV Sugar Men and Women	IV Sugar Women only	OLS Men and Women	OLS Men and Women	IV Sugar Men and Women	IV Sugar Women only
Sugar cultivation 1866	-0.004***	-0.005***	0.003	0.009	0.159***	0.105**	-0.577	-0.567
	(0.001)	(0.001)	(0.014)	(0.025)	(0.046)	(0.044)	(0.351)	(0.358)
Coffee cultivation 1863	-0.002	-0.001	-0.001	-0.000	0.143***	0.091**	0.074*	0.081*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.035)	(0.037)	(0.045)	(0.045)
Distance to railway 1884		-0.000***	-0.000	-0.000		-0.010***	-0.016***	-0.016***
		(0.000)	(0.000)	(0.000)		(0.002)	(0.004)	(0.004)
Distance to road 1884		0.001***	0.001***	0.001*		-0.045***	-0.053***	-0.055***
		(0.000)	(0.000)	(0.001)		(0.007)	(0.009)	(0.009)
Instrument controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	0.72	0.72	0.72	0.55	7.84	7.84	7.84	7.35
Observations	203,185	203,185	203,185	103,854	202,541	202,541	202,541	103,561
Number of clusters	423	423	423	423	423	423	423	423
R-squared	0.054	0.054	-	-	0.264	0.276	-	-

Notes: Each observation is an individual. The table reports the OLS as well as the 2nd stage estimates of an IV regression with sugar as endogenous variable. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Sectoral employment

Dependent Variable	Working in agricultural sector (Yes=1/No=0):				Working in manufacturing (Yes=1/No=0):			
	OLS Men and Women	OLS Men and Women	IV Sugar Men and Women	IV Sugar Women only	OLS Men and Women	OLS Men and Women	IV Sugar Men and Women	IV Sugar Women only
Sugar cultivation 1866	-0.026***	-0.021***	0.064	0.104*	0.014***	0.011***	-0.039	-0.063*
	(0.007)	(0.006)	(0.048)	(0.060)	(0.004)	(0.003)	(0.027)	(0.037)
Coffee cultivation 1863	-0.012**	-0.005	-0.004	-0.002	-0.002	-0.004*	-0.005*	-0.009**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.002)	(0.002)	(0.003)	(0.004)
Distance to railway 1884		0.001**	0.001***	0.002***		-0.001***	-0.001***	-0.001***
		(0.000)	(0.001)	(0.001)		(0.000)	(0.000)	(0.000)
Distance to road 1884		0.008***	0.008***	0.011***		-0.002***	-0.002***	-0.004***
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Instrument controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	0.32	0.32	0.32	0.32	0.15	0.15	0.14	0.17
Observations	137,031	137,031	137,031	53,404	137,031	137,031	137,031	53,404
Number of clusters	423	423	423	423	423	423	423	423
R-squared	0.193	0.210	-	-	0.044	0.048	-	-

Notes: Each observation is an individual. The table reports the OLS as well as the 2nd stage estimates of an IV regression with sugar as endogenous variable. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. The sectors the economy is divided in for the analysis are agriculture, manufacturing, services and other. in Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Female household heads and age at first marriage

Dependent variable	Household head female				Marriage age			
	OLS	OLS	IV Sugar	IV Sugar	OLS	OLS	IV Sugar	IV Sugar
Sugar cultivation 1866	-0.003***	-0.004***	0.004	0.007	0.111***	0.067**	-0.221	-0.253
	(0.001)	(0.001)	(0.008)	(0.008)	(0.032)	(0.032)	(0.276)	(0.259)
Coffee cultivation 1863	0.001	0.000	0.001	0.001	0.106***	0.065**	0.108***	0.057*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.028)	(0.028)	(0.029)	(0.031)
Distance to railway 1884		-0.000		0.000		-0.008***		-0.011***
		(0.000)		(0.000)		(0.002)		(0.003)
Distance to road 1884		-0.001***		-0.001***		-0.033***		-0.037***
		(0.000)		(0.000)		(0.006)		(0.007)
Instrument controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	0.14	0.14	0.14	0.14	19.35	19.35	19.35	19.35
Observations	87,632	87,632	87,632	87,632	99,985	99,985	99,985	99,985
Number of clusters	423	423	423	423	423	423	423	423
R-squared	0.067	0.067	-	-	0.132	0.139	-	-

Notes: The sample includes only women. The table reports the OLS as well as the 2nd stage estimates of an IV regression with sugar as endogenous variable. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

5.5 Discussion

In summary, we argue that our empirical analysis has demonstrated two broad tendencies: First, the sugar component of the Cultivation System caused a biased sex ratio (while the coffee component did not) which was often associated with a greater independence for women in the late colonial period. Second, after controlling for historical infrastructures, the intensity of sugar and coffee cultivation has no clear impact on contemporaneous female empowerment whereas distance to the major colonial roads and railways has a strong and mainly negative impact on education, work in modern sectors, and marriage age. Our results are consistent with a labor demand interpretation whereby the creation of a colonial infrastructure contributed to an industrialization process that opened up opportunities for women in the modern sectors.[11]

Many other papers in the literature have found a great degree of cultural persistence over time (Giuliano, 2017). In our view, our research is most strongly related to Grosjean and Khattar's (2015) analysis of the inflow of male convicts to colonial 18th- and 19th-century Australia (and the subsequent shock to the sex ratio) and its effect on gender balance. How can we explain that 19th century cultural norms appear to have been persistent among the male-dominated areas in Australia but not in our own Javanese sample where the shock to the sex ratio happened at almost the same time?

We argue that a key factor was most likely the very different size and character of the land areas and the resulting levels of population densities on the two islands. Differences in population densities between Java and Australia are to a great extent explained by the much higher land productivity on Java where a very efficient production of rice made it possible to support a substantially larger population than in the largely arid or semi-arid Australia. A higher population density, in combination with large infrastructural investments made necessary by the specific character of agricultural technology during the Cultivation System, both suggest that it was easier for Javans to interact with and migrate between other communities on the island than in Australia where several population concentrations in the country (for instance Perth in the northwest and Sydney in southeast) where relatively isolated from each other through most of its recent history. This greater mobility can explain at least in part why a historical shock to a particular area on Java did not persist through time to a similar extent as in Australia. Also, in absolute terms the bias of the historical sex ratio

[11] In spirit, this finding is similar to Dell and Olken (2017) who also find that the Cultivation System contributed importantly to Java's industrialization.

studied by Grosjean and Khattar is much larger than the bias we analyze in sugar cultivating districts on Java. Thus, if the historical effect on gender norms and labor outcomes was relatively small in our study area compared to the Australian case, there is also less scope for long-term persistence of the effect. Nevertheless, we think it is important and interesting to study relatively small biases of the sex ratio and their impacts, since such patterns occur probably in various places compared to the rather exceptional situation in Australia.

6. Concluding remarks

In this paper, we have analyzed the long-run implications of the Cultivation System on Java on female empowerment. Whereas historical sugar cultivation does not seem to have a long-term effect on female empowerment via culture, norms and preferences, it might have had through colonial infrastructure and the later industrialization

We believe that a promising area of future research would be to further explore the reasons why certain historical shocks have long-lasting effects whereas others do not. Such an analysis might also address the problem that studies finding cultural persistence are probably overrepresented in the literature since one might suspect a publication bias against studies showing that historical shocks actually had no lasting effects. In some sense, our study might contribute in this way by demonstrating an episode of non-persistent cultural effects of a major historical shock.

References

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos. 2011. "Marrying up: The Role of Sex Ratio in Assortative Matching." *American Economic Journal: Applied Economics* 3 (3):124–57.

Acemoglu, Daron, David H Autor, and David Lyle. 2004. "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy* 112 (3):497–551.

Acemoglu, Daron, Simon Johnson, and James Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review*.

Acemoglu, Daron, Simon Johnson, and James A Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *The Quarterly Journal of Economics* 117 (4):1231–94.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *The Quarterly Journal of Economics* 128 (2):469–530.

Algemeen Verslag betreffende de Koffijcultuur op Java. 1871. Departement van Kolonien.

Angrist, Josh. 2002. "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation." *The Quarterly Journal of Economics* 117 (3):997–1038.

Baten, Jörg, Mojgan Stegl, and Pierre van der Eng. 2013. "The Biological Standard of Living and Body Height in Colonial and Post-Colonial Indonesia, 1770–2000." *Journal of Bioeconomics*, 1–20.

Becker, Gary S. 1991. "A Treatise on the Family, Enl. Ed." *Cambridge, Mass: Harvard*.

Bobonis, Gustavo J, and Peter M Morrow. 2014. "Labor Coercion and the Accumulation of Human Capital." *Journal of Development Economics* 108:32–53.

Boomgaard, Peter. 1989. "Java's Agricultural Production 1775-1875." *Economic Growth in Indonesia* 1940:97–132.

Bosma, Ulbe. 2013. *The Sugar Plantation in India and Indonesia: Industrial Production, 1770-2010*. Cambridge University Press.

Bunn, Christian, Peter Läderach, Oriana Ovalle Rivera, and Dieter Kirschke. 2015. "A Bitter Cup: Climate Change Profile of Global Production of Arabica and Robusta Coffee." *Climatic Change* 129 (1–2):89–101.

Chiappori, Pierre-Andre, Bernard Fortin, and Guy Lacroix. 2002. "Marriage Market, Divorce Legislation, and Household Labor Supply." *Journal of Political Economy* 110 (1):37–72.

Clarence-Smith, William Gervase. 1994. "The Impact of Forced Coffee Cultivation on Java, 1805–1917." *Indonesia Circle* 22 (64):241–64.

Cribb, Robert B, and Rigobert Bonne. 2010. *Digital Atlas of Indonesian History*. NIAS Press Copenhagen, Denmark.

Dell, Melissa, and Benjamin A Olken. 2017. "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java." *Mimeo*.

Departement van Landbouw, Nijverheid en Handel. 1933. *Volkstelling 1930*. Batavia, Dutch East Indies.

Dippel, Christian, Avner Greif, and Dan Trefler. 2017. “Outside Options, Coercion, and Wages: Removing the Sugar Coating.”

Dippel, Christian, Avner Greif, and Daniel Trefler. 2015. “The Rents from Trade and Coercive Institutions: Removing the Sugar Coating.” National Bureau of Economic Research.

Donaldson, Dave. 2010. “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.” National Bureau of Economic Research.

Elson, Robert E. 1990. “Peasant Poverty and Prosperity under the Cultivation System in Java.” *Indonesian Economic History in the Dutch Colonial Era, Monograph Series35/Yale University Southeast Asia Studies, Yale Center for International and Area Studies, New Haven, Connecticut*.

Engerman, Stanley L, and Kenneth L Sokoloff. 2000. “History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World.” *The Journal of Economic Perspectives* 14 (3):217–32.

Estrada, Leonel Lara, Livia Rasche, and Uwe A Schneider. 2017. “Modeling Land Suitability for Coffea Arabica L. in Central America.” *Environmental Modelling & Software* 95:196–209.

Fasseur, C. 1992. *The Politics of Colonial Exploitation: Java, the Dutch, and the Cultivation System*. SEAP Publications.

Fernández, Raquel. 2013. “Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century.” *The American Economic Review* 103 (1):472–500.

Giuliano, Paola. 2017. “Gender: An Historical Perspective.” National Bureau of Economic Research.

Google maps. 2017. 2017. https://www.google.se/maps/place/Java/@-7.1668462,106.8576255,7.09z/data=!4m5!3m4!1s0x2e7aa07ed2323237:0x86fe1c59d6abed60!8m2!3d-7.6145292!4d110.7122465.

Grosjean, Pauline A, and Rose Khattar. 2015. “It’s Raining Men! Hallelujah?” *Mimeo*.

Hansen, Casper Worm, Peter Sandholt Jensen, and Christian Volmar Skovsgaard. 2015. “Modern Gender Roles and Agricultural History: The Neolithic Inheritance.” *Journal of Economic Growth* 20 (4):365–404.

Java Rail Map. 2017. 2017. http://flights.indonesiamatters.com/files/2011/10/2000px-Java_Transportation_Network.svg_.png.

Jayasinghe, PKSC, and Masao Yoshida. 2010. “Development of Two GIS-Based Modeling Frameworks to Identify Suitable Lands for Sugarcane Cultivation.” *Tropical Agriculture and Development* 54 (2):51–61.

Jedwab, Remi, and Alexander Moradi. 2016. “The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa.” *Review of Economics and Statistics* 98 (2):268–84.

Karadja, Mounir, and Erik Prawitz. 2016. “Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States.”

KITLV. 1884. “Kaart Aangevende: 5. de Spoor- / Stoomtram- / Post- En Groote Wegen.” KITLV Digital Image Library. http://media-kitlv.nl/all-media/indeling/detail/form/advanced/start/160?q_searchfield=java&f_uiterlijke_vorm%5B0%5D=Maps.

Koloniaal Verslag. 1868. ’s-Gravenhage : Algemeene Landsdrukkerij.

Michalopoulos, Stelios, Alireza Naghavi, and Giovanni Prarolo. 2014. “Trade and Geography in the Spread of Islam.” *The Economic Journal*.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge university press.

Nugent, Jeffrey B, and James A Robinson. 2010. “Are Factor Endowments Fate?” *Revista de Historia Economica-Journal of Iberian and Latin American Economic History* 28 (1):45–82.

Nunn, Nathan. 2008. “Slavery, Inequality, and Economic Development in the Americas.” *Institutions and Economic Performance*, 148–80.

Qian, Nancy. 2008. “Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance.” *The Quarterly Journal of Economics* 123 (3):1251–85.

Raffles, Thomas Stamford. 1830. “The History of Java, Vol. 2.”

Ricklefs, Merle Calvin. 2008. *A History of Modern Indonesia since c. 1200*. Palgrave Macmillan.

Sen, Amartya. 1990. “More than 100 Million Women Are Missing.” *The New York Review of Books* 37 (20):61–66.

Sutherland, Heather. 1979. *The Making of a Bureaucratic Elite: The Colonial Transformation of the Javanese Priyayi*. 2. Published for the Asian Studies Association of Australia by Heinemann Educational Books (Asia).

Teso, Edoardo. 2017. “The Long-Term Effect of Demographic Shocks on the Evolution of Gender Roles: Evidence from the Transatlantic Slave Trade.” *Journal of the European Economic Association* forthcoming.

Touwen, Jeroen. 2001. *Shipping and Trade in the Java Sea Region, 1870-1940: A Collection of Statistics on the Major Java Sea Ports*. KITLV Press.

Van Velthuizen, Harrij. 2007. *Mapping Biophysical Factors That Influence Agricultural Production and Rural Vulnerability*. Food & Agriculture Org.

Van Zanden, Jan Luiten. 2004. “On the Efficiency of Markets for Agricultural Products: Rice Prices and Capital Markets in Java, 1823–1853.” *The Journal of Economic History* 64 (4):1028–55.

———. 2010. “Colonial State Formation and Patterns of Economic Development in Java, 1800–1913.” *Economic History of Developing Regions* 25 (2):155–76.

Appendix

A.1. 1814 Analysis

Population data from 1814 comes from Raffles (1830). Here again, the names of the administrative units where matched with the names of the maps from 1857. A map showing the jurisdictions from 1814 is shown in map below. Reasons why we only have data for around half of the island in 1814 are, i) that not the whole island was under British control, ii) the name of the administrative units from 1814 could not be matched with the name of the jurisdictions in 1857 in the map of Cribbs. Reason i) is responsible for the missing data in central Java that was only colonized after the Java War in 1825-1830.

Figure A.1.1.: Map of 1814 jurisdictions

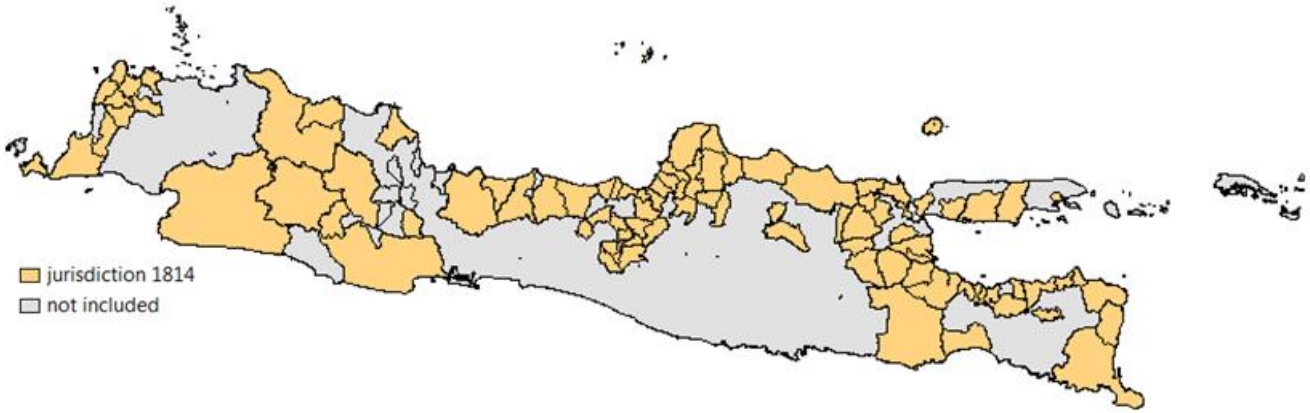


Table A.1.1.: Summary statistics of the 1814 data

Variables	**Obs.**	**Mean**	**St. Dev.**	**Min**	**Max**
Data set A: Pre-Cultivation System					
Area Covered: 1814 Subsample					
Sex Ratio 1814 (men/women)	86	0.99	0.08	0.82	1.35
Sex Ratio 1930 (men/women	86	0.95	0.04	0.83	1.08
Population Density 1814 (inhab. per sqkm)	86	113.18	154.18	0.82	1115.74
Sugar cultivation 1866 (percent of land area)	86	0.33	0.74	0	3.27
Coffee cultivation 1863 (percent of land area)	86	1.72	3.30	0	16.84
Geographical controls					
Sugar suitable land (share of land area)	86	0.66	0.39	0	1
Coffee suitable land (share of land area)	86	0.20	0.29	0	1
Distance to port (in km)	86	76.02	78.89	3.49	324.15

Table A.1.2.: Historical Sex ratio and sugar cultivation, 1814 jurisdictions only

Dependent Variable	Sex Ratio 1814		Sex Ratio 1930	
	OLS	OLS	OLS	OLS
Sugar cultivation 1866	-0.001	-0.001	-0.004	-0.004
	(0.011)	(0.013)	(0.006)	(0.006)
Coffee cultivation 1863		0.001		0.002
		(0.003)		(0.002)
Constant	0.961***	0.961***	0.970***	0.970***
	(0.044)	(0.044)	(0.023)	(0.023)
Mean dep. variable	0.99	0.99	0.95	0.95
Mean sugar cultivation	0.43	0.43	0.43	0.43
Mean coffee cultivation	1.73	1.73	1.73	1.73
Distance to Port	Yes	Yes	Yes	Yes
Suitability variables	Yes	Yes	Yes	Yes
Population Dens. 1814	Yes	Yes	Yes	Yes
Observations	86	86	86	86
R-squared	0.060	0.061	0.165	0.174

Notes: Level of aggregation: Jurisidicts from 1814, Study area: Jurisdictions of 1814 only. Rrobust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix A.2. Additional figures and tables

Figure A.2.1. Main ports in Central and Eastern Java

Note: The black dots on the map show the nine ports that were used to calculate the distance to from the geographical center of each historical district. Starting from the left, the ports are: Cirebon, Tegal, Pekalongan, Semarang, Surabaya, Pasuruan, Probolinggo, Sitobondo and Banyuwangi.

Figure A.2.2. Railways and roads on Java and Madura in 1884

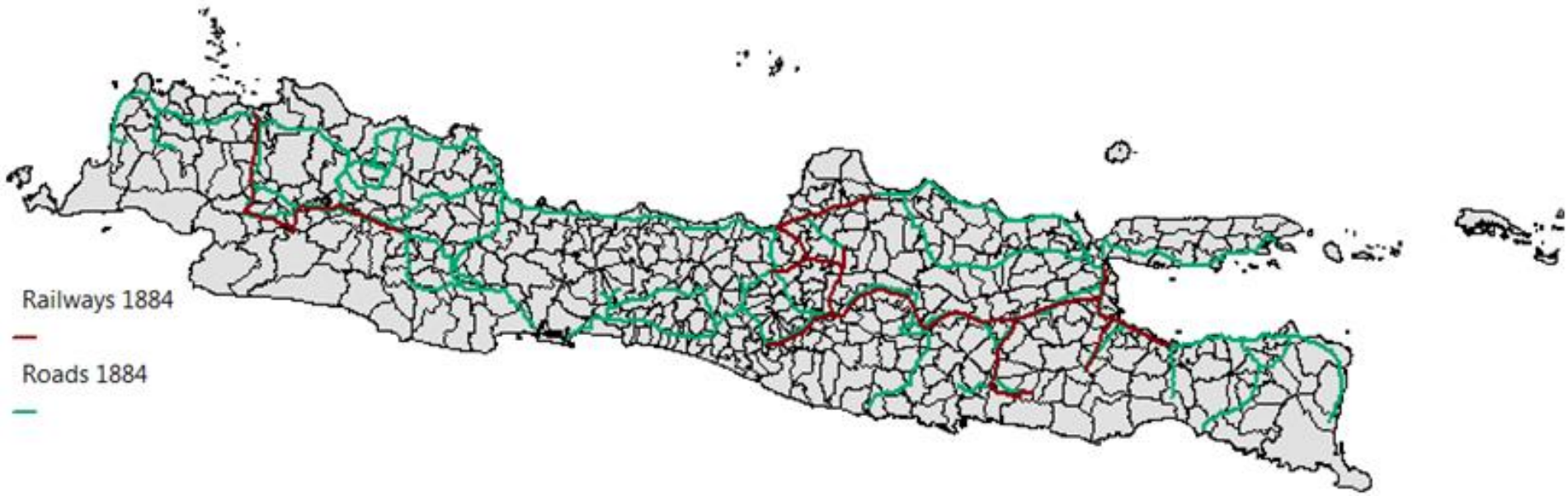
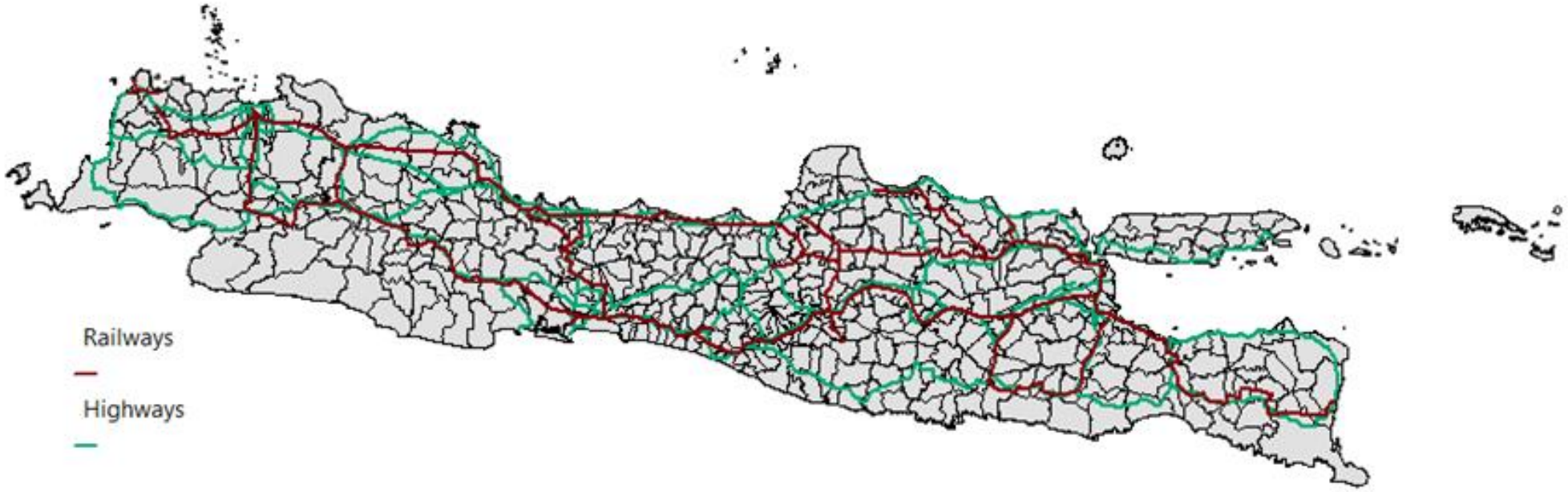


Figure A.2.3. Main railways and highways in 2017



Note: Sources: For highways: Google Maps (2017) and for railways: Java Rail Map (2017).

Figure A.2.4. Scatter plots of sugar suitability and distance to port with sugar cultivation in 1866

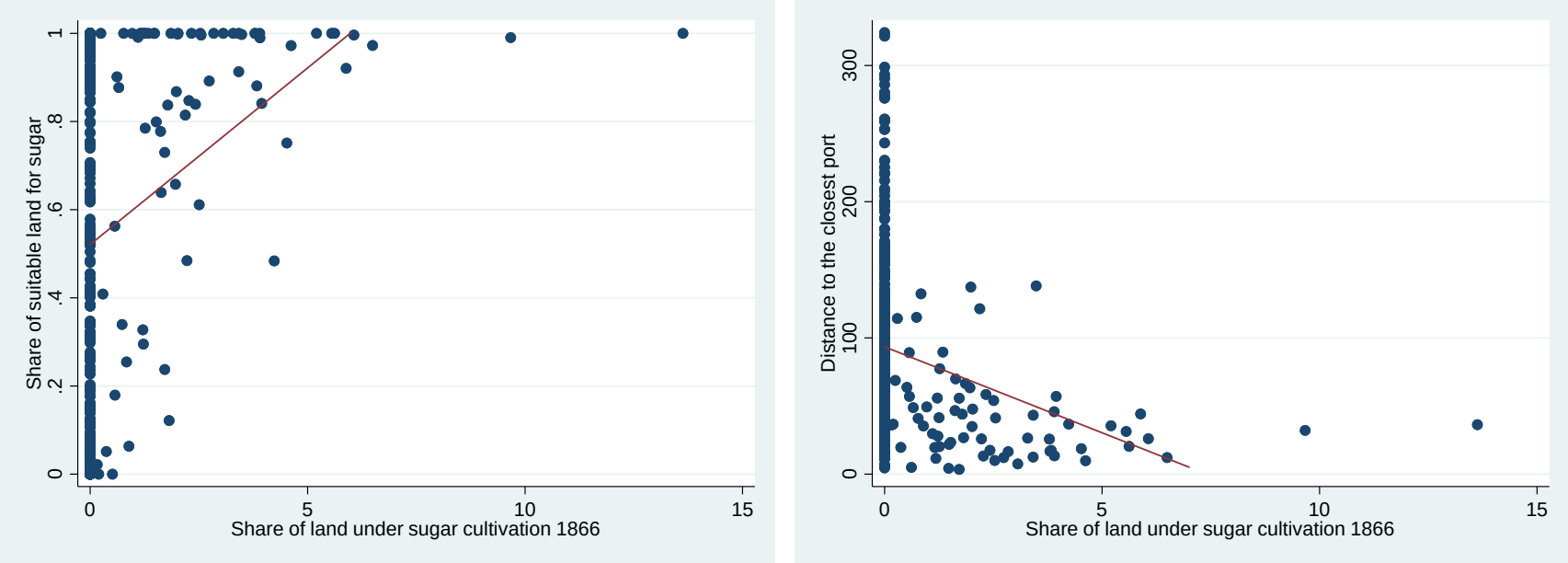


Figure A.2.5. Scatter plots of coffee suitability and distance to port with coffee cultivation in 1863

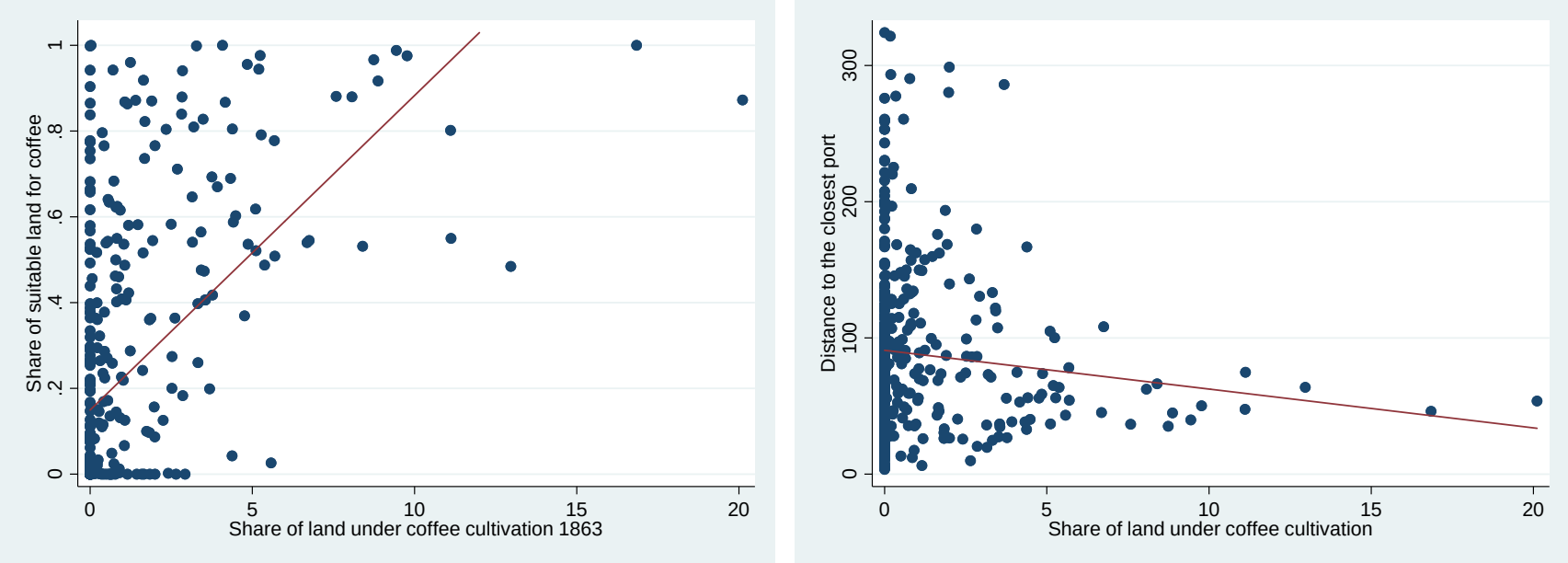


Table A.2.1: Summary Statistics, Regency level

Variables	Obs.	Mean	St. Dev.	Min	Max
Data set B: Regency level					
Area Covered: Java and Madura, level of aggregation: regency					
1930 census variables					
Divorced females (share)	87	0.05	0.02	0.02	0.09
Household head female (share)	87	0.13	0.03	0.08	0.20
Married females (share)	87	0.43	0.03	0.39	0.61
Sex ratio (men/women)	87	0.95	0.03	0.86	1.07
Variables from colonial reports					
Sugar cultivation 1866 (percent of land area)	87	0.08	0.24	0	1.48
Coffee cultivation 1863 (percent of land area)	87	0.88	1.59	0	9.63
Geographical controls					
Sugar suitable land (share of land area)	87	0.57	0.34	0	1
Coffee suitable land (share of land area)	87	0.20	0.22	0	0.89
Distance to Port (in km)	87	86.72	58.58	15.36	304.83

Table A.2.2.: Sex ratio 1930 and crop cultivation during the Cultivation System, Coffe cultivation instrumented

Dependent variable	Sex ratio 1930			
	IV Coffee			
Coffee cultivation 1863	0.006**	0.005**	0.005*	0.005*
	(0.003)	(0.002)	(0.003)	(0.003)
Sugar cultivation 1866		-0.004***		-0.003**
		(0.001)		(0.001)
Constant	0.972***	0.975***	0.875***	0.887***
	(0.007)	(0.007)	(0.042)	(0.042)
Mean dep. variable	0.96	0.96	0.96	0.96
Mean sugar cultivation	0.42	0.42	0.42	0.42
Mean coffee cultivation	0.98	0.98	0.98	0.98
Distance to port	Yes	Yes	Yes	Yes
Suitability variables	Yes	Yes	Yes	Yes
Add. geo. controls	No	No	Yes	Yes
Historical controls	No	No	Yes	Yes
Observations	425	425	425	425

Notes: Each observation is a historical district. The excluded instrument in the first stage is the interaction term between *Coffee suitability* and *Distance to Port*. Additional geographical controls include distance to the coast, elevation, slope and general soil constraints. Historical controls include population density in 1930 and the years of colonial rule. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2.3.: Instrument relevance

	Sugar cultivation 1866		Coffee cultivation 1863	
Dependent variable	Men and Women	Women only	Men and Women	Women only
	OLS	OLS	OLS	OLS
Distance to port	-0.000	-0.000	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Sugar suitability	0.790**	0.770**		
	(0.383)	(0.387)		
Distance to port*	**-0.009*** **	**-0.009*** **		
Sugar suitability	**(0.003)**	**(0.003)**		
Coffee suitability			5.367***	5.371***
			(1.385)	(1.381)
Distance to port*			**-0.039*** **	**-0.039*** **
Coffee suitability			**(0.010)**	**(0.010)**
Mean dep. variable	0.42	0.42	0.97	0.97
Instrument controls	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Observations	203,185	103,854	203,185	103,854
F Statistic (Instruments, bold coefficients)	10.12	9.85	16.48	16.59
R-squared	0.227	0.224	0.447	0.448

Notes: Each observation is a historical district. The excludable instrument is the interaction terms between *Distance to port* and *Sugar suitability*. Additional geographical controls include distance to the coast, elevation, slope and general soil constraints. Historical controls include population density in 1930 and the years of colonial rule in 1930. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2.4: Sectoral employment in services and remaining sectors

Dependent Variable	Working in service sector (Yes=1/No=0):				Working in other sectors (Yes=1/No=0):			
	OLS Men and Women	OLS Men and Women	IV Sugar Men and Women	IV Sugar Women only	OLS Men and Women	OLS Men and Women	IV Sugar Men and Women	IV Sugar Women only
Sugar cultivation 1866	0.011**	0.009*	-0.018	-0.032	0.001	0.000	-0.006	-0.009
	(0.005)	(0.005)	(0.034)	(0.043)	(0.001)	(0.001)	(0.007)	(0.006)
Coffee cultivation 1863	0.013**	0.009*	0.009	0.010*	0.001	0.000	-0.000	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)
Distance to railway 1884		-0.000	-0.000	-0.000		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Distance to road 1884		-0.005***	-0.005***	-0.006***		-0.000**	-0.001**	-0.000
		(0.001)	(0.001)	(0.001)		(0.000)	(0.000)	(0.000)
Instrument controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	0.46	0.46	0.46	0.46	0.08	0.08	0.08	0.08
Observations	137,031	137,031	137,031	53,404	137,031	137,031	137,031	53,404
Number of clusters	423	423	423	423	423	423	423	423
R-squared	0.098	0.103	-	-	0.005	0.005	-	-

Notes: Each observation is an individual. The table reports the OLS as well as the 2nd stage estimates of an IV regression with sugar as endogenous variable. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. The sectors the economy is divided in for the analysis are agriculture, manufacturing, services and other. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2.5: Instrumental regressions without infrastructure variables

Dependent Variable	IV Sugar Men and Women					
			Working in:			
	Labor Force Participation	Years of education	Agriculture	Manufacturing	Services	Other
Sugar cultivation 1866	0.010	-0.525	0.078	-0.034	-0.037	-0.007
	(0.016)	(0.389)	(0.060)	(0.026)	(0.041)	(0.007)
Coffee cultivation 1863	-0.002	0.147***	-0.012**	-0.001	0.013**	0.001
	(0.002)	(0.039)	(0.006)	(0.003)	(0.006)	(0.001)
Distance to railway 1884						
Distance to road 1884						
Instrument controls	Yes	Yes	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. variable	0.46	0.46	0.46	0.46	0.46	0.08
Observations	203,185	202,541	137,031	137,031	137,031	137,031
Number of clusters	423	423	423	423	423	423

Notes: Each observation is an individual. The table reports the OLS as well as the 2nd stage estimates of an IV regression with sugar as endogenous variable. The excluded instrument in the first stage is the interaction term between *Sugar suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. The sectors the economy is divided in for the analysis are agriculture, manufacturing, services and other. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2.6: Labor market participation and education level

Dependent Variable	Labor Force Participation: (Yes=1/No=0)		Years of education	
	IV Coffee Men and Women	IV Coffee Women only	IV Coffee Men and Women	IV Coffee Women only
Sugar cultivation 1866	-0.004***	-0.006***	0.107**	0.108**
	(0.001)	(0.002)	(0.044)	(0.045)
Coffee cultivation 1863	0.001	0.005	0.146	0.134
	(0.006)	(0.009)	(0.118)	(0.119)
Distance to railway 1884	-0.000**	-0.000**	-0.009***	-0.010***
	(0.000)	(0.000)	(0.002)	(0.003)
Distance to road 1884	0.001***	0.001**	-0.043***	-0.046***
	(0.000)	(0.001)	(0.008)	(0.008)
Instrument controls	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Mean dep. variable	0.72	0.55	7.84	7.35
Observations	203,185	103,854	203,185	103,561
Number of clusters	423	423	423	423

Notes: Each observation is an individual. The table reports the 2nd stage estimates of an IV regression with coffee as endogenous variable. The excluded instrument in the first stage is the interaction term between *Coffee suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. The sectors the economy is divided in for the analysis are agriculture, manufacturing, services and other. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2.7: Sectoral employment

Dependent Variable	Working in agricultural sector (Yes=1/No=0):		Working in manufacturing (Yes=1/No=0):	
	IV Coffee Men and Women	IV Coffee Women only	IV Coffee Men and Women	IV Coffee Women only
Sugar cultivation 1866	-0.021***	-0.022***	0.011***	0.008*
	(0.006)	(0.006)	(0.003)	(0.004)
Coffee cultivation 1863	-0.007	-0.015	0.003	0.002
	(0.015)	(0.016)	(0.009)	(0.011)
Distance to railway 1884	0.001*	0.000	-0.000***	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)
Distance to road 1884	0.008***	0.009***	-0.002***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Instrument controls	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Mean dep. variable	0.32	0.32	0.14	0.17
Observations	137,031	53,404	137,031	53,404
Number of clusters	423	423	423	423

Notes: Each observation is an individual. The table reports the 2nd stage estimates of an IV regression with coffee as endogenous variable. The excluded instrument in the first stage is the interaction term between *Coffee suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. The sectors the economy is divided in for the analysis are agriculture, manufacturing, services and other. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2.8: Sectoral employment

Dependent Variable	Working in service sector (Yes=1/No=0):		Working in other sectors (Yes=1/No=0):	
	IV Coffee Men and Women	IV Coffee Women only	IV Coffee Men and Women	IV Coffee Women only
Sugar cultivation 1866	0.009*	0.013**	0.000	0.000
	(0.005)	(0.005)	(0.001)	(0.001)
Coffee cultivation 1863	0.004	0.010	0.000	0.002
	(0.016)	(0.017)	(0.002)	(0.002)
Distance to railway 1884	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Distance to road 1884	-0.005***	-0.006***	-0.000**	0.000
	(0.001)	(0.001)	(0.000)	(0.000)
Instrument controls	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Mean dep. variable	0.46	0.46	0.08	0.08
Observations	137,031	53,404	137,031	53,404
Number of clusters	423	423	423	423

Notes: Each observation is an individual. The table reports the 2nd stage estimates of an IV regression with coffee as endogenous variable. The excluded instrument in the first stage is the interaction term between *Coffee suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. The sectors the economy is divided in for the analysis are agriculture, manufacturing, services and other. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2.9.: Female household heads and age at first marriage

Dependent variable	Household head female		Marriage age	
	IV Coffee	IV Coffee	IV Coffee	IV Coffee
Sugar cultivation 1866	-0.003***	-0.004***	0.112***	0.066**
	(0.001)	(0.001)	(0.032)	(0.032)
Coffee cultivation 1863	-0.003	-0.003	0.038	0.041
	(0.004)	(0.004)	(0.116)	(0.098)
Distance to railway 1884		-0.000*		-0.008***
		(0.000)		(0.002)
Distance to road 1884		-0.001***		-0.034***
		(0.000)		(0.007)
Instrument controls	Yes	Yes	Yes	Yes
Add. geo controls	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes
Contemp. controls	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Mean dep. variable	0.14	0.14	19.35	19.35
Observations	87,632	87,632	99,985	99,985
Number of clusters	423	423	423	423

Notes: The sample only includes women. The table reports the 2nd stage estimates of an IV regression with coffee as endogenous variable. The excluded instrument in the first stage is the interaction term between *Coffee suitability* and *Distance to Port*. Instrument controls included at the district level are the share of suitable land for sugar and coffee and distance to the closest port. The vector of additional geographical controls includes the variables distance to the coastline, elevation, slope, general soil constraints. Further, the historical variables population density in 1930 and years under colonial rule in 1930 are included and as contemporaneous controls on the district level the population density in 2008 and sex ratio in 2008. Individual controls included are age and age squared. Clustered standard errors at the historical district in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.2.10.: Bins used to create sugar suitability index

Sugar suitability Bins:	**(4)**	**(3)**	**(2)**	**(1)**	**(2)**	**(3)**	**(4)**
Temperature (C)	<19	20-22	23-25	26-30			
Rainfall (mm)			<1200	1200-1500	1500- 2000	2000-2500	>2500
Elevation (m)			<0	0- 100	100-500	500- 1000	>1000
Slope (%)				<15	15-30	30-60	>60
Soil ph	<5	5.0-5.5	5.5-6.0	6.0-7.0	7.0-7.5	7.5-8.5	>8.5

Notes: For Temperature, Rainfall and Soil ph bin categories are based on Dippel et al. (2017) and Jayasinghe and Yoshida (2010). For elevation and slope, the bins were adjusted to fit the special caractersitics of Java and Madura.

Table A.2.11.: Bins used to create coffee suitability index

Coffee suitability Bins:	**(4)**	**(3)**	**(2)**	**(1)**	**(2)**	**(3)**	**(4)**
Temperature (C)	<10	10-16.5	16.5-18	18-22	22-24.5	24.5-31	>31
Rainfall (mm)	<1000	1000-1300	1300-1575	1575-1900	1900-3000	>3000	
Elevation (m)	<0	0-300	300-475	475-1300	>1300		
Slope (%)				<40	40-60	>60	
Soil ph	<4	4.0-5.0	5.0-5.35	5.35-6.35	6.35-6.5	6.5-8.0	>8.0

Notes: For Temperature, Rainfall, Soil ph and Slope, bin categories are based on values presented in Estrada et al. (2017). For Elevation, values are based on Bunn et al. (2015).

Table A.2.12.: Data sources to create suitability indices

Data sources for suitability indices	
Temperature	
Variable description:	Average temperature in degree Celsius per year for the period 1961-1990
Source:	*Fick, S.E. and R.J. Hijmans, 2017. Worldclim 2: New 1-km spatial resolution climate*
Link:	http://worldclim.org/version2
Rainfall	
Variable description:	Average rainfall in mm per year for the period 1961-1990
Source:	*Fick, S.E. and R.J. Hijmans, 2017. Worldclim 2: New 1-km spatial resolution climate*
Link:	http://worldclim.org/version2
Elevation	
Variable description:	Altitude above sea level in meter
Source:	*Jarvis, A., H.I. Reuter, A. Nelson, E. Guevara, 2008, Hole-filled SRTM for the globe Version 4*
Link:	http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1
Slope	
Variable description:	Slope in %, calculated in ArcGIS based on the elevation data
Source:	*Jarvis, A., H.I. Reuter, A. Nelson, E. Guevara, 2008, Hole-filled SRTM for the globe Version 4*
Link:	http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1
Soill Ph	
Variable description:	Measures the acidity and alkalinity of the soil
Source:	*Fao, Iiasa, and Isscas Isric. "JRC: Harmonized World Soil Database (version 1.2)." FAO, Rome, Italy and IIASA, Laxenburg, Austria (2012)*
Link:	http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/