CER-ETH - Center of Economic Research at ETH Zurich

Did History Breed Inequality? Colonial Factor Endowments and Modern Income Distribution

Matthew J. Baker, Christa N. Brunnschweiler, and Erwin H. Bulte

Working Paper 08/86
June 2008

Economics Working Paper Series

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Did History Breed Inequality? Colonial Factor Endowments and Modern Income Distribution

Matthew J. Baker[1], *Christa N. Brunnschweiler*[2], and *Erwin H. Bulte*[3]

[1] Hunter College, City University of New York
Email: Matthew.Baker@hunter.cuny.edu

[2] CER-ETH Center of Economic Research at ETH Zurich, Switzerland
Email: cbrunnschweiler@ethz.ch

[3] Development Economics Group, Wageningen University and Department of Economics, Tilburg University
Email: Erwin.bulte@wur.nl

June 12, 2008

Abstract: We explore the relation between historical population density in former colonies and modern income distribution. A theoretical model highlights the potentially opposing effects of native population density on incentives for colonists to conquer or settle in new territories. While an abundant supply of native labor is an "asset" that drives up land rents, it is also a "liability" that makes land acquisition by colonists more difficult and reduces returns to peacable migration. Conflicts over land, sowing the seeds for inequality by creating a landed élite living off rents, are especially likely to emerge for intermediate native population densities. Results are confirmed by detailed empirical tests highlighting the curvilinear relationship between native population density and modern income inequality. Finally, using population density as an instrument for inequality in the former colonies, we demonstrate that there is no causal relationship running from income distribution to economic growth.

JEL Codes: O15, N30, N50

Keywords: inequality, growth, factor endowments, population density, conflict, colonization

Acknowledgements: Erwin Bulte thanks the Dutch Organisation for Scientific Research (N.W.O.), grant nr. 452-04-333 for financial support.

1. Introduction

One of the classical questions in development economics concerns the nature of the two-way relationship between income inequality and economic development. The research tradition dates back at least to the 1950s (e.g., Lewis 1954, Kuznets 1955, Kaldor 1956), but is still relevant today (e.g., Anand and Kanbur 1993a,b, Persson and Tabellini 1994, Alesina and Rodrik 1994, Clarke 1995, Forbes 2000, Easterly 2007). One of the key issues in this literature is whether inequality is a barrier to growth and development. This could be the case if inequality triggers growth-impeding redistributive policies, or if it undermines institutional quality (resulting in political instability) and negatively affects the incentives and ability of a large part of the population to invest in various forms of capital (e.g., Easterly 2007). The exact nature of the relationship between inequality and growth remains a subject of debate, as there is also evidence suggesting that growth is promoted if income is concentrated in the hands of an élite with a relatively large propensity to invest (see Forbes 2000 for recent evidence along these lines).

The debate on the relation between inequality and economic performance naturally leads to another fundamental question regarding the origins of inequality itself. Recent research emphasizes that the distribution of income in countries is not invariant with respect to underlying characteristics of these countries. In other words, inequality is an endogenous variable, and jointly determined along with other macroeconomic variables of interest. A plausible case can be made that inequality is closely linked to the institutional structure of economies. However, this in turn begs an answer on the origins of institutions. The social conflict view on institutions argues that institutional structures are persistent and evolve only slowly in response to contemporary pressures (see Glaeser et al. 2004, Bourguignon and Verdier 2000, Acemoglu *et al*. 2001, Engerman and Sokoloff 1997, 2002 and Sokoloff and Engermann 2000, but also the discussion below). If so, we would expect inequality to change slowly, as well, and we could trace the roots of current inequality patterns back to factors dating from much earlier periods in human history.

In this paper we explore the main determinants of inequality in former colonies. Following seminal work on incentives faced by colonial settlers several centuries ago (Engermann and Sokoloff 1997, 2002, Sokoloff and Engermann 2000, Acemoglu *et al*. 2001, 2002), we examine the impact of historical factor endowments on current economic outcomes. However, whereas earlier work often focused on the effects on modern income levels and the distribution of world income and has been dominated by qualitative and quantitative studies, we focus on the modern distribution of income within countries and develop a theoretical model that is able to explain the empirical pattern of development.

The factor endowment of particular interest to us is native population density in pre-colonial days. It has been suggested that densely populated areas were an attractive "prize" for predatory colonists, as they were relatively prosperous (and therefore offered richer spoils), and because abundant native labor would prove useful as an input in plantation agriculture, mining, or other activities. This in turn led these areas to be characterized by extractive economic and social structures (e.g. Engermann and Sokoloff 1997, 2002, Sokoloff and Engermann 2000, Acemoglu *et al*. 2002). However, and in contrast, more numerous populations most likely were also better capable of defending themselves, and therefore harder to conquer and subdue. If so, the relation between native population density and historical patterns of colonization, institution-building, and economic development could be curvilinear, with limited colonization and subjugation effort at both extremes of the distribution (i.e. in very thinly and very densely populated native societies). With institutional persistence, this would eventually affect contemporary social structures and institutional design, in turn influencing modern income distributions within countries.

The main objective of this paper is twofold. First and foremost, we analyze the relationship between population density in former colonies, incentives for colonizers, and resulting patterns of inequality within the confines of a simple general equilibrium trade model. In our minds, this is an important addition to the literature, which is largely atheoretical. The model we develop is stylized and simplified in the sense that we ignore forced labor (slaves) and social hierarchies (i.e. rulers, or authoritarian elites). Following previous literature, the model proceeds from the assumption that inequality is to a large degree the result of the segmentation of the population in a given colony into a colonial elite and a common, native class. The model treats both struggles over natural resource endowments (i.e., land) and peaceable migration in a general equilibrium framework of colonial activity and trade.

The most interesting result generated by the theory is that the relationship between subsequent inequality (as measured by either a land gini coefficient or an income gini coefficient) and initial population density is non-monotonic. In fact, the degree of inequality is likely increasing in initial population density for low population densities, and decreasing in population density at higher population densities. Our theory also generates predictions about the nature of production and subsequent inequality. For example, greater land intensity of production increases subsequent inequality in the colony.

We test our hypothesis against some of the other popular explanations on the origins of inequality. The chain of events that we envisage is a three-step process: (i) the impact of historical population density on settlement and labor allocation decisions (i.e. the social and economic structures) of colonial settlers, (ii) the impact of historical social and economic structures on modern-day institutions, and (iii) the subsequent impact on the distribution of income in modern societies. However, for lack of detailed information on historical social structures and institutions, we follow Acemoglu *et al*. (2001, 2002) and leave the second link implicit in the empirical analysis. We then use the results of the abovementioned analysis as the starting point for the development of a novel instrument for income inequality, and use this instrument to analyse the relation between inequality and income levels and growth. We find no strong causal relationship running from inequality to contemporary income levels or growth in former colonies.

The paper is organized as follows. In section 2 we briefly summarize previous research on the colonial origins of economic development and formulate a new hypothesis on the interaction of population density and technology, which we support with historical evidence. Section 3 provides a formal model to illustrate the main economic mechanisms driving colonization and subsequent social and economic development paths. In section 4 we empirically test our theory for former colonies and find that the results are consistent with the theory, highlighting the curvilinear relation between native population density in pre-colonial times and modern inequality patterns. Moreover, our hypothesis proves robust to controlling for other possible factors that influenced development in former colonies, such as geography and mineral abundance. We then use historical population density as an instrument for current inequality and explore the relation between inequality and income as well as income growth. Section 5 concludes.

2. A Hypothesis of Colonization Patterns and Inequality

2.1 Economic theories and a novel hypothesis

Engerman and Sokoloff (henceforth ES) have proposed that the roots of modern-day wealth disparities in former New World colonies lie in differing factor endowments at the time of colonization. They hypothesized that three factors were crucial for development: the climate and soil type; the presence of precious metals; and the density of the native population. The structural differences related to initial factor endowments then tended to persist, "not only because certain fundamental characteristics of New World economies were difficult to change, but also because government policies and other institutions tended to reproduce them" (ES 2000: 223).[1]

The first factor – which we will term "geography" – determined whether the colony was suitable for growing high-value crops such as sugarcane and other plantation-type products. If this was the case, the society which emerged was generally characterized by a stark class division between the landowning, white colonist élite, and the rest of the population, both native and of African slave descent. Conversely, if geographic conditions lent themselves more to the production of grains and hays, with limited economies of scale, a more egalitarian social structure typically emerged, based on small landholdings.

Mineral abundance – the second factor – greatly influenced the early development of Spanish American countries such as Mexico and Peru. The Spanish authorities awarded large land-grants (*encomiendas*) to the narrow European-descent élite, often including claims to mines as well as to the native labor force residing near the estate. The main focus of the conquerors was the exploitation of the mineral resources – primarily silver and gold – and the extraction of labor and tribute from the native population; low-value "bulk" primary products such as livestock remained under-exploited until the advent of cheaper transport in the 19th century (Waites 1999).[2] Again, areas rich in mineral resources tended to generate economic structures with very unequal wealth distributions between the owners of the mines and the mine workers.

The third factor, the native population density, influenced all New World colonies in a similar fashion, and according to ES, its interaction with the other factors was crucial for

[1] The idea that institutions persist for a long time can be found already in Wittfogel's (1957) analysis of China, Russia and the Ottoman Empire. Further evidence supporting the view that in particular colonial institutions and economic structures have persisted is provided in Young (1994), La Porta et al. (1999), Coatsworth (1999), Mariscal and Sokoloff (2000), and North et al. (2000).

[2] The main natural resources exported from Spanish America to the Old World until the 19th century were precious metals. It is estimated that the Americas – foremost Mexico and Peru – produced about 102'000 tons of silver from the time of their discovery until 1810, equivalent to 85% of world production. Similarly, the New World also delivered around 70% of the world's gold produced during that period (Waites 1999: 35-36).

determining the development patterns in the colonies. Sparsely inhabited territories encouraged settler-type colonization, based on immigrant laborers and small-scale farmers of European descent. The result was the emergence of relatively homogeneous populations with similar wealth and human capital distributions. Conversely, large native populations provided abundant cheap labor to work on the colonists' estates, plantations and mines.[3] Consequently, substantial native populations at the time of colonization are linked to more unequal societies.

Though not uncontroversial, the historically-founded research of ES has inspired several recent studies; and although the focus of ES was on the Americas, their approach has been extended to encompass all former colonies, or even the entire world. For example, Easterly (2007) uses the ratio of soil suitability for growing wheat versus sugarcane as an instrument for inequality. He then finds that inequality predicts development, measured alternatively by present-day income levels, schooling rates, and institutional quality. Acemoglu *et al*. (2001) use settler mortality as an instrument for institutional quality to explain modern income levels. They find that areas with a more hospitable climate and disease environment are linked to higher levels of institutional development, and to higher present-day income levels (see also Easterly and Levine 2003, Rodrik et al. 2004). Acemoglu *et al*. (2002) use early native urbanization rates and population densities to explain their hypothesis of a "reversal of fortune". Because densely populated areas were more likely to be subjected to extractive institutions by European colonizers, they ended up with institutional set-ups which were less favorable to broad investment activities of the type needed to later take advantage of the industrial revolution. Densely populated – and therefore rich – countries at the time of colonization consequently have lower income levels today, and vice versa.

[3] An interesting case of how the conquerors used the more numerous native population to their own advantage is given by the (forced) relocation of Tlascaltecan families into northern Mexico starting in 1580. The Tlascaltecans—Spanish allies since the days of Cortés—were intended to protect the sparse white settlers against the "rude tribes" along the borders; in return for this "service", they benefited from certain privileges such as the exemption from tribute (Bolton and Marshall 1920: 59).

Previous studies assumed that the relationship between economic and institutional development and early factor endowments is more or less a linear one, not only for the New World colonies examined by ES, but also for colonies (and even non-colonies) elsewhere. However, this assumption turns out to be false for one fundamental factor, namely native population density at the time of colonization. Figure 1 plots the influence of population density in 1500 ca. (in logs) on modern inequality, measured by the Gini index around 1960, for all former colonies.[4] Clearly, the relationship is *not* linear. Instead, it roughly follows an inverted-U shape. It also appears from the years of colonization in Figure 1(b) that the change in colonization pattern – where population density is linked to the maximum observed inequality levels – occurred around the start of the 18th century.[5] What determined this fundamental change in colonization patterns, which appears to have influenced economic development to this day?

<< *Insert Figure 1 about here* >>

We develop and test the hypothesis that European colonization, and the subsequent distribution of wealth in the colonies persisting until modern times, is attributable to the ratio of land to native inhabitants which the early colonizers were confronted with, and to the (military) technology of which they disposed. Colonizers traded off the benefits of a higher native population against the cost of military conflict and subjugation of said population. We hypothesize that (i) very low native population densities tended to lead to settler-type colonization, with egalitarian social and political structures and hence more equal wealth distributions (as in ES); at the other end of the spectrum, (ii) very high native population densities discouraged forced change as this would have been too costly. Colonization is delayed

[4] Scatter plots using the top quintile income share as the measure of inequality show the same pattern. Note that here as in the rest of the paper, we do not include those countries which were colonized for a brief period in the 20th century when colonial empires were already on the wane, such as Libya (a colony from 1911-1947) and Ethiopia (1936-1941).The data is discussed in more detail in Section 4.

[5] Two countries—South Africa and Indonesia—are listed with the dates of the official founding of the first permanent trading post in the early 17th century. In fact, though it is difficult to assign an exact date to the moment where most of the present-day area was conquered, the inland territorial expansion remained circumscribed until the 18th century.

until permitted by military innovations; and when eventually implemented, the colonizers often took over the administrative centre previously held by the native ruling class, leaving the rest of the social and institutional structure largely unaltered. Depending on existing conditions, this could often imply a relatively more equal distribution of wealth. In between lie those colonies where (iii) native population density was relatively high, but still low enough to permit conquest and widespread control by force. The result was the emergence of societies with a large income gap between the colonizing élite and the rest of the population.[6]

2.2 Historical background

Most former colonies – with the notable exception of Australia and New Zealand, whose discovery in the late 18th century can indeed be regarded as more coincidental – were already familiar to Europeans by the end of the 17th century, either through first-hand discovery or second-hand information. They had known the countries along the North African coast of the Mediterranean for centuries, and had heard of the sub-Saharan African civilizations and the far-off Asian countries from exchanges with merchants and adventurers.[7] It would be difficult to contend that there was nothing to tempt conquerors into Africa or Asia, given the presence of precious metals and other coveted natural resources such as spices. Yet the vast territorial conquests in the New World during the 15th and 16th centuries were not echoed by similar advances in Africa and Asia: there was to be no more major territorial conquest until well into the 18th century (not counting the small and sparsely populated island colonies of Mauritius and the Seychelles, taken over by the French in the first half of the 18th century).

---

[6] Our hypothesis not only helps to explain the colonization pattern of those countries that were ultimately conquered, but may also shed light on why some countries defied colonization, most notably the densely populated and technologically advanced China and Japan. Military conquest failed repeatedly, and European powers were limited to competing for their commercial positions in these countries.

[7] Portuguese traders began exploring the African coast southward from the 15th century onwards. Vasco da Gama and Albuquerque rounded the Cape of Good Hope, sailed up the East African coast and reached as far as India and Malacca at the turn of the 16th century, plundering port towns and setting up new trading posts along the way. During the same period, Columbus, Cortés and Pizarro conquered (most of) the Americas; and soon afterwards, the Spaniards established the westward route from Central America to their new Pacific colony of the Philippines.

In their early conquests, the Europeans' military technology brought them a definite advantage over the native populations, which not only lacked firearms but also had few defensible bases to fall back on. However, the relatively easy victories over the great civilizations of the Aztecs and the Incas in the Americas, and a few decades later in the Philippines, could not readily be repeated in other parts of the world, where the obstacles facing the small groups of Europeans were more formidable (see e.g. Guilmartin 1995, Parker 1996). European conquest in Africa was halted by heavy native resistance and the hostile disease environment, evidenced for example during the failed Portuguese attempt to establish a colony in Luanda in the 1570s (Waites 1999), or their abortive expedition up the Zambesi river in the 1690s (Black 1995, Keim 1995). In Asia, the Europeans found themselves facing opponents at roughly equal technological levels. Their main early advantage lay in the gunned sailing ship, which enabled them to terrorize coastal towns;[8] however, this technology was soon imitated and put to effective use, as the naval battles off the Malabar Coast in 1503, off Tunmen on the Chinese coast in 1522, and in the sultanate of Aceh demonstrate (Guilmartin 1995, Parker 1996, Waites 1999). Moreover, the early warships carried too few troops to successfully launch amphibious attacks on the relatively heavily populated interior (Cable 1998, Waites 1999).

In short, in both Africa and Asia – with the exception of the Philippines – native opposition limited European territorial expansion to the occupation of coastal forts and sometimes the immediate *hinterland*. Where forceful conquest proved too costly to accomplish, Europeans engaged in more peaceful and spatially limited activities, either farming in order to replenish the supplies of the merchant vessels (as in the Cape Colony), or specializing in trade – mainly in gold and slaves in West Africa, and in spices and textiles in Asia. They "undertook military expenditure either to coerce reluctant buyers or in order to safeguard themselves

[8] For more on the historical importance and the advances in early naval warfare, see Glete (2005).

against attack from their European rivals; the cost of defence would otherwise have eaten up all their trading profits" (Parker 1996: 132).[9]

From the mid-18th century, however, the Europeans were able to rapidly extend their colonial dominion inland into hitherto impregnable regions, starting with major advances in the Indian subcontinent and Indochina (see e.g. Parker 1996) and culminating in the "scramble for Africa".[10] This shift in the balance of power between "the West and the rest" coincides with the completion of a remarkable series of technological innovations in Europe, which historians have called the "military revolution". The term, originally coined by Michael Roberts in 1955, encompasses the introduction of three key innovations between 1560-1660: the capital ship with its broadside guns and greater sea-endurance; the development of gunpowder weapons "as the arbiter of battles and sieges"; and in response to the latter, the development of the artillery fortress or *trace italienne* (Parker 1996: 159).[11] These innovations enabled the Europeans to overthrow – and maintain – many of those densely populated and relatively advanced areas which had previously defied conquest.[12]

Hence, population density and technology notably influenced European colonization patterns. We aim to complete the picture by drawing the link to the social and economic

[9] This is the period of the rise of the great European Trading Companies, foremost the Dutch and British East India Companies. Although territorial conquest was not among their chief objectives, it gradually did come as a consequence of efforts to preserve their trading positions (Waites 1999). Where there was organized resistance, "the conqueror [negotiated] with it in order to break it all the better later" when the opportunity arose, most commonly by seeking allies among the opponent's rivals (Ferro 1997: 32; for more on the importance of native collaboration for imperialism, see Robinson 1972). This strategy proved effective for the British in India, the Dutch in Indonesia, and later the French in the Maghreb.

[10] This paper does not aim at exploring the reasons behind the scramble for Africa in the 19th century; for more on this subject, see e.g. Louis (1976) and Förster et al. (1988). We simply contend that population density and European technology played a role in determining colonization patterns and economic development paths in former colonies.

[11] Gunpowder weapons already came into use in the 14th century; but early firearms such as the *arquebus* were notoriously inaccurate and difficult to handle, and it wasn't until the 16th century that they gradually became more effective on land and onboard ships, and as small firearms (Parker 1996).

[12] A good example is given by India: the early colonial strategy of the British (then represented by the British East India Company) was to concentrate their trade efforts in relatively small and weak areas with few European competitors, such as Golconda, the Carnatic, and parts of Bengal. But by the end of the 18th century, with the help of native *sepoy* armies trained by Europeans and armed with European-made weapons, the British were able to successfully challenge the smaller independent states, as well as their French rivals on the subcontinent. From the rest of Bengal, they moved on to Bihar and Orissa, before their progress was halted by the Marathas, who had in the meantime adopted European military techniques. The whole of the Indian subcontinent was finally conquered in the 19th century (see Parker 1996, Waites 1999).

structures which emerged, and explain the inverse-U shaped relationship with inequality. As regards the colonies conquered before the 18th century—the Americas as well as the Philippines—our hypothesis closely follows that of ES. Areas with low native population densities were relatively easily conquered; they encouraged settler-type colonization and tended to produce egalitarian societies. More substantial native populations were aggressively subdued, and previous systems of forced labor and tribute were partly borrowed and adapted to create highly unequal "plunder economies". Where native death rates were particularly high due to disease, enslavement, or hunger, slaves and *mestizos* provided the cheap labor force necessary to work the plantations and mines.

By the end of the 17th century, the colonizers seem to have reached the limits of the native population size which they could overthrow and extract cheap labor from by force. But as the Europeans' technological level increased, some new territories followed the pattern of the more densely populated Latin American colonies. For example, the early population densities and present-day inequality levels in Southern Africa are comparable to the more populous New World colonies such as Guyana, Cuba and Panama. However, native resistance to the early European colonists was decisive enough to delay full conquest of the southernmost part of Africa until the 18th century. But once imposed, colonial rule gave rise to similar development paths as those seen in densely populated American countries: white settler colonists competed for land, labor and influence with the native Africans. Typically, strict racial restrictions in all economic and social areas emerged, with a small number of white landowners and capitalists forming the privileged élite (Gellar 1995).

Although central institutions may have been relatively easy to capture thanks to their technological superiority, imposing colonial authority in vast new territories with ever-larger native populations proved much more difficult. While colonizers were able to retain the technological edge, it was not the overwhelming superiority as in the Central and South

American colonies, and completely subduing the new colonies proved to be very costly. As the Europeans conquered more and more populous areas, their strategy leaned towards adapting the native social and economic structure without imposing profound changes or adding to the existing inequalities by creating a new elite.[13] In sum, the essential feature of European colonial rule in Africa and Asia is the conservation of the traditional social and economic institutions, "while the administrative apex was monopolized by a white élite" (Waites 1999: 147).

3. A Decentralized Colonial Model of Inequality

While there has been a wealth of empirical work on initial factor endowments and how endowments influence the subsequent performance of the economy, to date there has been little theoretical work, in spite of the seemingly obvious nature of the costs and benefits to various types of colonial activities to colonial powers. We develop a simple general equilibrium model of the colonization process to see how one might expect colonists to behave when confronted with different sorts of factor endowments in a colony. The basic operating principle of the model is that the way in which colonists deploy resources in a colony are driven by the relative

[13] In Africa, relatively populous Nigeria was not conquered until the 1880s, and administration and taxation in the colony once conquered were only possible with the support of the Muslim Fulani aristocracy. But in order to gain this, the British colonialists—who had in fact already espoused the humanitarian and abolitionist causes in the late 18th century—had to reach a compromise with the slave-owning Fulani, which meant that the traditional slave society was only phased out by the 1930s (Waites 1999). Most of the Nigerian population was left out of the colonial economic system, which essentially precluded economic development (Gellar 1995); these early differences seem to be reflected in modern income inequality levels, with a relatively high Gini index of 51. Uganda is an interesting example of a very densely populated colony with present-day inequality levels comparable to those of Canada and Australia. Slave raids had been frequent in this region, and European protection in the 19th century was welcomed (Waites 1999). Rural farmers were not displaced by colonial policies, and even enjoyed some increased economic opportunities producing cash crops such as cocoa and coffee for European export. The colonial impact on the social and economic structure in Uganda was limited to the central administration by a small group of white colonists (Gellar 1995). India had early native population densities at the upper bounds of all former colonies (25.81 per square km on average); yet its modern income inequality level is close to that of the U.S. and Uruguay (35.56 vs. 36.02 and 36.61, respectively). The British secured authority over the Indian subcontinent by a mix of aggression and collaboration with the traditional élites. When the British Government officially took over the rule of the colony from the East India Company in 1774, it adapted and even reinforced the traditional social system. Above all in the countryside where most Indians lived (and continue to live), change was hardly evident and modernization practically wholly absent. The colonial rulers "chose security over development" as they lacked the power resources to effectively transform and develop the entire society (Waites 1999: 184).

marginal benefits of different types of activities.[14] This allows us to characterize the nature of colonial activities in equilibrium, and develop explicit measures of inequality for the colony in terms of initial conditions, including resources endowments, native population, and the nature of production in the colony.

The model is a standard general equilibrium trade model. The only departure from the basic trade model is the inclusion of a measure of the potential colony's capacity for self defense. The model allows assessment of the nature and degree of colonization and the resulting degree of inequality strictly as functions of initial conditions in the potential colony. Initial conditions include the resource endowment, the population, and the production technique of the potential colony, in addition to the production technique favored by the colonist in the potential colony.

The ultimate goal of the exercise is to explore the relationship between native population density in the colony, and settlement decisions of the colonizers. The latter affect inequality and institutional development, which in turn (following Acemoglu *et al*. 2001) we postulate may affect contemporary inequality patterns.

3.1 Model Setup

We begin by setting up a basic model in which there is one potential colony and a colonial power. In section 3.4, we expand this model to include a spectrum of potential colonies, which enables us to relate the results of this section to the historical evolution of colonialism.

There are three goods: X, a manufactured good produced only in the colonizing country (perhaps because of some specific factor of production such as capital there); Y, a good the colonizing country wishes to consume but is produced only in the potential colony (perhaps a raw material); and H, a non-traded good produced only in the colony.

[14] We do not explicitly include a (colonial) government to the model, nor do we aim to capture collective decision-making by the natives. Such political economy concerns are left for future research—here we aim to analyze the decentralized equilibrium when colonists and natives interact (violently through conflict, or peacefully through rental arrangements or trade).

The inhabitants of the colonizing country consume goods *Y* and *X*, while inhabitants of the potential colony consume goods *X* and *H*. The typical inhabitant of the colonial power has utility function:

$$u_c = x^{1-\alpha} y^{\alpha} . \tag{1}$$

The utility function for inhabitants of the potential colony is:

$$u_n = h^{\delta} x^{1-\delta} .\text{[15]} \tag{2}$$

Let $I_i, i = c, n$ represent aggregate income in each of the two locations (throughout, variables pertaining to the colonizer are subscripted with *c* and those pertaining to the potential colony are subscripted with *n*). Let $p_j, j = X, Y, H$ represent the prices of each of the goods. Maximization of (1) and (2) subject to budget constraints results in demand functions for goods *X*, *Y*, and *H*:

$$X^d = X_c^d + X_n^d = \frac{(1-\alpha)I_e}{p_x} + \frac{(1-\delta)I_n}{p_x},$$

$$Y^d = \frac{\alpha I_c}{p_y}, \qquad H^d = \frac{\delta I_n}{p_h} \tag{3}$$

Equilibrium prices, incomes, and factor allocations can be obtained by jointly solving the demand equations (3), coupled with a description of resource endowments, production technologies, and equilibrium conditions.

To specify factor endowments, suppose that the potential colonizer has a total endowment of labor, L_c. This endowment may be applied to three things: production of good *X* at home; production of good *Y* in the colony (which requires peaceful migration to the colony); and conquest in the colony. Let L_x denote labor allocated to production of good *X* (at home), L_{cy} devote labor allocated to production in the colony by the colonist (peaceful migration), and

[15] If one is uncomfortable with utility functions, (1) and (2) could just as easily be interpreted as composite production functions through which a composite consumption good is produced.

let M denote labor devoted to conquest in the colony. Labor devoted to conquest will ultimately determine colonists' land endowment in the colony. The basis for these three types of tasks is that we want to compare peaceful colonization with violent colonization directed towards redistribution of productive resources.

Simplifying matters, we assume that peaceful migrants to the colony (L_{cy}) only produce good Y. Production-minded colonists are incapable of producing the non-tradable good H (or unwilling to do so). This assumption might be motivated because of specific cultural knowledge unique to the native population in production of H. Collecting this information, the endowment constraint faced by the colonizer is:

$$L_x + L_{cy} + M = L_c \tag{4}$$

The inhabitants of the potential colony have some total endowment of labor L_n and some endowment of land, A_n. These two factors of production can be allocated between producing good H and good Y, so we have the following resource constraints in the colony:

$$A_{ny} + A_{nh} = A_n; \qquad L_{ny} + L_{nh} = L_n \tag{5}$$

In the potential colony goods Y and H are produced according to standard Cobb-Douglas production technologies:

$$Y = L_y^{1-\beta_y} A_y^{\beta_y}, \qquad H = L_h^{1-\beta_h} A_h^{\beta_h} \tag{6}$$

Thus, $\beta_i, i = Y, H$ in (6) measures the land intensity of production in each of the two goods, while $1-\beta_i$ measures the labor intensity of production. For good X, produced solely at home by the potential colonizer, we assume that production occurs via the simple linear production function:

$$X = \theta_x L_x \tag{7}$$

The specification is the simplest possible way of characterizing the opportunity costs of engaging in colonial activities, as each unit of labor devoted to colonial activities results in θ_x

units of foregone output at home. One could make this specification more complex by adding some fixed factors such as land, but this would not change results.

The potential colony is also endowed with a capacity to defend itself, which we will denote as *D*, and the final ingredient in the basic structure of the model is a description as to how the quantities *M* and *D* interact. Describing this interaction allows us to characterize how land endowments are split across the native and colonizing populations, and admits a description of the returns to engaging in conquest relative to peaceable migration to the colonizer. We follow Hirschleifer (1995) and Grossman and Kim (1995) and employ a standard contest function to specify the amount of land seized by colonists. If the total endowment of land in the colony is given by *A*, then the amount of land seized by colonists is:

$$A_c = A\frac{M}{M+D}. \tag{8}$$

It follows that land remaining in the hands of the native population is given by:

$$A_n = A - A_c = A\frac{D}{M+D}. \tag{9}$$

Obviously $A_n = A$ for $M = 0$. If colonists engage in more land-grabbing behavior, the immediate result is less land for the native population. Note that since by assumption colonizers cannot produce good *H*, it follows that $A_c = A_{cy}$, or all land possessed by colonizers is devoted strictly to the production of good *Y*.

The model is closed through description of equilibrium conditions. In the potential colony, factors of production are allocated across production of *Y* and *H* so as to equate marginal value products:

$$MVP_{Ay} = MVP_{Ah}; \qquad MVP_{Ly} = MVP_{Lh} \tag{10}$$

The marginal value product functions in (10) may have corner solutions, where, for example, $MVP_{Ay} < MVP_{Ah}$. Intuitively, this might occur because colonists have acquired so much land from the native population that natives prefer to put all of their land into use in the production of good *H*. As we discuss below, solutions to the conditions (10), taking as given colonial activity in the colony, are typically of the form:

$$A_{ny}^{*} = \xi_0 A_n - \xi_1 A_{cy}, \qquad L_{ny}^{*} = \psi_0 L_n - \psi_1 L_{cy}, \tag{11}$$

where $\xi_0, \xi_1, \psi_0, \psi_1$ are positive constants. The factor allocations in (11) indicate that as colonists acquire more land, the natural impact of this is native substitution of land or labor away from production of good *Y*. This effect is more severe for land than it is for labor. As colonists acquire more land, the endowment of land of the native population is directly reduced. Not only does the native population in the colony wish to allocate less land to production of *Y*, but the native population has less land in total to allocate.

While one can certainly postulate more complex motives on the part of colonists (we discuss some extensions below), we imagine that the colonists allocate their resources between home production of good *X*, peaceful migration for direct production of good *Y*, and conquest so that marginal value products across the three activities are equalized. This requires that the marginal value product of labor employed in production of *X* at home be equal to the marginal value product of labor in production of good *Y* in the colony. That is:

$$MVP_{Ly} = MVP_{Lx}. \tag{12}$$

Equilbrium conquest of land requires that:

$$MVP_{Ay} a_c = MVP_{Lx}. \tag{13}$$

In (13), a_c is amount of land per unit of labor devoted to conquest. Expression (8) gives us an expression for *total* land acquired, so if acquired land is equally divided among all labor devoted to conquest, we find that a_c in equation (13) becomes:

$$a_c = \frac{1}{M} A \frac{M}{M+D} = \frac{A}{M+D}. \tag{14}$$

The model is now fully specified, and can be solved using endowment constraints (4) and (5); production technologies given by (6), (7), and (8); demand functions (9); and the equilibrium conditions (10), (12), and (13). The solution method we employed is described fully in appendix 1, which shows how the system can be boiled down, after eliminating wages and prices, to a four-equation system solvable for $M, L_{ny}, L_{cy}, A_{ny}$.

The solutions are (under the working assumption that all variables have an interior solution):

$$M^* = \alpha\beta_y L_c - D + \delta\beta_h (D + L_n)\,, \qquad (15)$$

$$L^*_{ny} = L_n - \delta(1-\beta_h)(L_n + D) \qquad (16)$$

$$L^*_{cy} = \alpha(1-\beta_y)L_c - L_n + \delta(1-\beta_h)(D + L_n) \qquad (17)$$

$$A^*_{ny} = A\frac{D - \delta\beta_h(D + L_n)}{\alpha\beta_y L_c + \delta\beta_h(D + L_n)}\,. \qquad (18)$$

One can see from inspection of (15-18) that there is a possibility of corner solutions where the optimal values are zero. We shall delve into these possibilities momentarily, and subsequent assumptions will rule out the possibility that L^*_{ny} and A^*_{ny} are zero. Before doing that, we draw attention to a key ambiguity introduced in (15). On the one hand, more formidable defenses imply reduced conquest effort by the colonizers ($dM^*/dD<0$). This follows simply from the contest function – given a certain choice of conflict effort, more formidable defenses reduce the amount of land that may be taken. But on the other hand, a larger productive native population implies enhanced conquest effort ($dM^*/dL_n>0$). This follows from the fact that increasing the number of natives bids up land rents. Depending on assumptions with respect to key parameters, the impact of native population (density) on the inflow of land grabbing colonizers is ambiguous, as greater population density is likely to both increase the size of the native labor force and the capacity of the potential colony to defend itself from conquest.

Given the above (interior) solutions, we can find expressions for total national income in the colony, the share of national income in the colony possessed by natives, and the share of the native population in the ultimate population of the colony. We form an expression for total national income in the colony using a sum-of-factor-payments approach:

$$I_n = w_h L_h + w_y L_y + r_h A_h + r_y A_y \tag{19}$$

Equating wages with marginal value products and using equilibrium values (15-18) in the resulting expression for (19) – expressions for necessary terms such as equilibrium prices and marginal products appear in appendix 1 – we find that total national income in the colony at an interior solution to (15-18) is given by:

$$I_n = \theta_x(\delta(D + L_n) + \alpha L_c) \tag{20}$$

The amount of income held by the native population in the colony is:

$$I_{nn} = \theta_x(D + L_n)\,. \tag{21}$$

Hence, the share of income held by natives in the colony is obtained by dividing (20) by (21):

$$s_{In} = \frac{D + L_n}{\delta(D + L_n) + \alpha L_c}\,. \tag{22}$$

Equation (22) says that the share of national income that the native population will be able to secure depends upon the size of the native aggregate labor endowment as well as the native capacity to defend themselves relative to the size of the colonial power (as measured by L_c). Further, the share of income held by natives depends upon the relative demand intensities of the two countries. Larger values of α, measuring the intensity of preferences among colonizers for good *Y*, reduce the equilibrium share of income held by natives. One can see from (15) and (17) how larger values of α translate into an increased presence in the colony by the colonizing country. Natives also wind up with a lower share of income if their preferences for the manufactured good (as measured by δ) are large. In these cases, natives are hurt because their demand for foreign goods results in lower prices for colonial goods.

Substituting M^* in (15) into (9) gives a direct expression for the ultimate share of land or resources retained by the native population:

$$s_{An} = \frac{D}{\alpha\beta_y L_c + \delta\beta_h (D + L_n)} \tag{23}$$

From (23), one can see some of the same effects underlying (22), with some additional insights. Again, the share of land retained by natives is increasing in their capacity to defend themselves, as measured by *D*, but decreasing in the total indigenous labor supply of the colony. The share of land held by natives is also decreasing in demand parameters in similar fashion to the income share in (22), but we can also see that the native land share falls if production of either good *Y* or good *H* is more land intensive.

Since we have treated all members of the native population and all colonists symmetrically, the only source of inequality in the model is between colonial "élites" and native inhabitants of the colony. One can therefore compute Gini coefficients as the difference between the population share held by natives and the income (or land) share held by natives. To develop a formal expression for the Gini coefficient, let *N* represent the population of the potential colony so that we can write the share of natives in the ultimate population of the colony as follows:

$$f_n = \frac{N}{N + L_{cy}^* + M^*} \tag{24}$$

Plugging the requisite values from (15) and (17) into (24), we find that:

$$f_n = \frac{N}{\alpha L_c + N - (1-\delta)(L_n + D)}. \tag{25}$$

From (25), one can see how a higher initial native population has both a direct and an indirect effect on the native population share. If *N* increases, this exerts a direct impact on (25) simply because *N* is explicitly part of the population share. However, one would also expect that a larger *N* also translates into a larger existing supply of labor in the colony and also an expanded capacity for defense, raising both D and L_n, which, according to (25), also increases the native population share. Income and land Gini coefficients, can be formed using (25) and (22) or (23). The income gini coefficient is given by:

$$g_I = f_n - s_{In}; \tag{26}$$

While the land gini coefficient is given by:

$$g_A = f_n - s_{An} \tag{27}$$

We now explore how the Gini coefficients in (26) and (27) systematically vary with conditions in the colony.

3.2 Inequality and Factor Endowments in Potential Colonies

Labor endowments and defensive capabilities should depend directly on the population of the native economy. Here, we assume this to be the case explicitly: let $D = D(N)$, $D' > 0$ and $L_n = L_n(N)$, $L_n' > 0$. It is also useful to consider the sum of these expressions to be the composite function $Z(N) = D(N) + L(N)$; from prevous assumptions, it follows $Z' > 0$. Plugging this specification into the expressions for the income gini coefficient (26) and simplifying results in:

$$g_I(N) = \frac{[N - Z(N)][\alpha L_c - Z(N)(1-\delta)]}{[\alpha L_c + N - (1-\delta)Z(N)][\delta Z(N) + \alpha L_c]} \tag{28}$$

Over the relevant range of population values, the expression in the numerator of (28) is approximately quadratic. To be more concrete about this, focus on the linear case, for which $Z(N) = zN$, and z is sufficiently small.[16] Equation (28) becomes:

$$g_I(N) = \frac{(1-z)N[\alpha L_c - (1-\delta)zN]}{[\alpha L_c + N - (1-\delta)zN][\delta zN + \alpha L_c]}. \tag{29}$$

The numerator in (28) is a quadratic function. Thus, we expect that the Gini initially increases in population density, but eventually begins to decrease. In terms of the historical record,

[16] Typically, it should be assumed that z is <1. This assumption conveys both a military and a technological advantage to the potential colonizer.

income inequality is predicted to be largest among those colonies with intermediate population densities.

The logic behind this result is quite simple. For a degree of inequality to be present in the colony, it is required that both the colonizers and the natives be present in sufficient numbers. For very small native population densities, incentives to migrate – either for peacable reasons or for reasons of conquest – are large. The result is a very large colonial presence relative to the native population. While natives in this situation wind up with very small shares of income or land, their population shares are also very small and the result is a rather insignificant dent in the Lorenz curve.

The opposite is true in situations in which the native population is very large. The large population deters both peacable migration and migration for conquest from the colonial power. While the resulting migration produces colonists who earn greater income than the native population, once again, the resulting population share of colonials is too small to greatly influence the Lorenz curve associated with the gini coefficients. Only for intermediate native populations is the resulting inflow of colonists large enough to comprise a large share of the colonial economy, but not so large that the share of the native population is unimportant.

To further explore the implications of the model, assume that $\beta_h = \beta_y = 1/2$, retain the assumption that $Z(N) = zN, z < 1$, and let $D(N) = Z(N)/2$ and $L_n(N) = Z(N)/2$. This means that each additional unit of population contributes to defense and labor supply equally. The main result of these assumptions is that we now have (from equations (15 and 17) $M^* = L^*_{cy}$, and as a result, we find that the total number of colonists (equally split between the two activities) is as follows:

$$C^* = M^* + L^*_{cy} = \alpha L_c - (1-\delta)zN . \quad (30)$$

From equation (28), if the colony has a population greater than $N = [(1-\delta)z]^{-1}\alpha L_c$, there will be no colonization. Put differently, a potential colony with a high enough population density will not be colonized, either peaceably or through the use of force. While this parameterization is of course highly specialized, it does lend insight into the basic predictions of the model. Consider Figure 2, which graphs as functions of initial population density, (i) the income Gini coefficient, (ii) total colonization effort in the colony (C^*), and (iii) the position of the point of inflection of the Lorenz curve. This diagram serves to illustrate the argument preceding equation (30) about the initial population conditions under which the largest degree of inequality is expected to be observed. One can see from figure 2 how the degree of colonial migration to the colony depends on initial population, how this maps into an inflection point on the Lorenz curve, and what this means for the ultimate value of both the Income Gini coefficient and the Land Gini coefficient, which, for this parameterization, are the same (compare equations (22) and (23)).

3.3 Income Inequality, Land Inequality, and the Nature of Production

In the previous section, we assumed $\beta_h = \beta_y = 1/2$. However, if we allow for differences in the land intensity of production across the home-produced good and the export good in the potential colony, we shall see that the land Gini can be either greater than or less than the income Gini. Particularly, in situations in which β_y is large, it is likely that the land Gini coefficient is much larger than the income Gini. This is important because it echoes the argument of many growth researchers, who seem to think that scale economies and land intensity in production influence the resulting nature of inequality.

As a first step, note the two optimal military and peaceable intervention equations (15) and (17). If it is the case that $\beta_y > 1/2$, the intercept term for M^* is larger than that of L_{cy}^*. The result is that the M^* function typically lies above L_{cy}^*. In fact, if one again maintains the assumptions that $Z = zN, z < 1$, $D = Z/2$, $L_n = Z/2$, one can again see from (15) and (17) that there are initial colonial populations sufficient to completely deter conquest or peaceable intervention. We have:

$$M^* = 0 \to N \geq \frac{\alpha\beta_y L_c}{z\left(\frac{1}{2} - \delta\beta_h\right)}, \qquad L_{cy}^* = 0 \to N \geq \frac{\alpha(1-\beta_y)L_c}{z\left(\frac{1}{2} - \delta(1-\beta_h)\right)}. \tag{31}$$

At either of these two points, whichever comes first, the Gini coefficient will undergo a discrete change in form. From (31) one can check that if $\frac{1}{2} > \beta_y(1-\delta) + \delta\beta_h$, peaceable migration ends at a lower value of initial population than military intervention, and, past the critical value of N for which $L_{cy}^* = 0$ given in (31), optimal military intervention becomes:

$$M_{Lcy=0}^* = \frac{\alpha(\delta\beta_h + (1-\delta)\beta_y)L_c - (1-\delta)D}{1-\delta(1-\beta_h)}. \tag{32}$$

The resulting income share is:

$$s_{In,L_{cy=0}} = \frac{\alpha(1-\beta_y)L_c + D}{\alpha(1-\delta(\beta_y - \beta_h))L_c + \delta D}, \tag{33}$$

while the resulting population and land shares, are, respectively::

$$s_{pn,L_{cy=0}} = \frac{N(1-\delta(1-\beta_h))}{N(1-\delta(1-\beta_h)) - D(1-\delta) + \alpha L_c(\beta_y - (\beta_y - \beta_h)\delta)}. \tag{34}$$

$$s_{An,L_{cy=0}} = \frac{(1-\delta(1-\beta_h))D}{(1-\delta(1-\beta_h))D + \alpha(\delta\beta_h + (1-\delta)\beta_y)L_c}. \tag{35}$$

In the alternative case, where $\frac{1}{2} \leq \beta_y(1-\delta) + \delta\beta_h$, we have $M^* = 0$ first, and this gives equilibrium peaceable migration as:

$$L_{cy,M=0}^* = \frac{\alpha(1-\beta_y + \delta(\beta_y - \beta_h))L_c - (1-\delta)L_n}{1-\delta\beta_h}, \tag{36}$$

The income share of the native population is:

$$s_{In,L_{cy=0}} = \frac{\alpha\beta_y L_c + L_n}{\alpha L_c(1+\delta(\beta_y - \beta_h)) + \delta L_n}, \tag{37}$$

and the native share of the ultimate colonial population is:

$$s_{pn,L_{cy=0}} = \frac{(1-\delta\beta_h)N}{(1-\delta\beta_h)N - (1-\delta)L_n + \alpha L_c(1-\beta_y + (\beta_y - \beta_h)\delta)}. \tag{38}$$

What does this mean for the underlying nature of equilibrium? Figure 3 plots two alternative cases: one in which the parameters are such that military intervention ceases first, and one in which peaceable migration ceases first. For the first case, we see that the land Gini reflects a greater degree of inequality than the income Gini. From these diagrams, one can see that more land intensive production likely results in a larger value of the Land Gini Coefficient.

3.4 A Quasi-Historical Argument

Historically, it appears that the first colonies generally had lower initial population densities than later colonies, and given our arguments, this should allow some direct analogies to be made between the time at which a potential colony was first colonized and the resulting degree of inequality. In this section, we discuss how one might modify our theoretical model to make some conjectures about the way in which colonization played out over history and as both the military and productive technologies of colonial powers increased. In section 4, of the paper, we explore the empirical implications of thes results.

We now imagine that instead of a single potential colony, there is a continuous spectrum of potential colonies indexed by $i \in [0,1]$. We suppose that the distribution of potential colonies is uniform, and that potential colonies are arranged in ascending order of population, and in particular, the population of the *i*th colony is $N(i) = \eta i$. As in previous sections, we assume that defensive capacities and labor supplies in the colonies is proportional to population: $Z(i) = zN(i), z < 1$, $D(i) = Z(i)/2$, $L_n(i) = Z(i)/2$. We make the additional simplifying assumption that $\beta_h = \beta_y = \frac{1}{2}$ across all potential colonies. These assumptions ensure that colonial efforts within any given colony will be equally divided between conquest and production.

We assume each colony now produces a differentiated product which enters into the colonial power utility function symmetrically. The utility function is now given by:

$$\ln U = (1-\alpha)\ln X + \alpha \int_0^1 \ln Y(i)di \tag{39}$$

Whereas in any given colony, the utility function is given by (2). The rest of the model is essentially set up as before and the solution procedure is the same. However, in addition to solving for equilibrium colonial intervention in each colony, we must now find both a reservation colony – i.e., the value of the index i^* above which no colonial effort will be exerted. This task takes a fair amount of algebra, but is not conceptually difficult. We outline the solution of the model in this situation in some detail in appendix 2. The chief results are as follows:

The "cutoff" potential colony i^* is:

$$i^* = \frac{\alpha L_c}{(1-\delta)z\eta} \tag{40}$$

The factor shares of the native population in colony $i \le i^*$ are:

$$\frac{L_{cy}(i)}{L_y(i)} = \frac{A_{cy}(i)}{A_y(i)} = \frac{i}{i^*}. \tag{41}$$

The share of income and land coincide and are given by:

$$s_{In}(i) = \frac{i}{\delta i + (1-\delta)i^*} \tag{42}$$

The native population shares for colonies $i \le i^*$ are:

$$s_{pn}(i) = \frac{i}{(i^* - i)(1-\delta)z + i}. \tag{43}$$

Combining (42) and (43) results in the Gini coefficient:

$$g_I(i) = g_A(i) = \frac{i(i^* - i)(1-z)(1-\delta)}{[\delta i + (1-\delta)i^*][(i^* - i)(1-\delta)z + i]} \quad i \le i^* \tag{44}$$

The expression(s) for the Gini coefficient in (44) reflect the same overall shape as that described in section 3.2, however, now the interpretation is a bit different. Rather than describing comparative statics of inequality with respect to initial population density, equation (44), together with equation (40), describe the spectrum of potential colonies that are likely to be colonized along with the resulting degree of income and/or land inequality. One can see from inspection of (44) that, for those colonies for which colonial intervention actually occurs, the degree of inequality bears an inverse-U shaped relationship with the value of initial population (which is directly proportional to the value of the index i by assumption.

To cast the results in terms of history as simply as possible, consider a two period model in which some technological progress occurs in the colonial power. In terms of the model, this situation might be reflected by a smaller value of z in the second period than in the first, reflecting the idea that in relative terms, the technology in the colony is now less advanced. From equation (40), this improvement in technology should increase the set of potential colonies that are actually colonized. From equation (44), the resulting inequality both before and after the technological advance can be deduced. One might then develop an understanding as to how and why the date at which colonial intervention began might be capable of explaining subsequent inequality.

Figure 4 illustrates the idea, graphing results for two different values of z,where z_1 is the relative technology prior to technological advancement in the colonial power, and z_2 is the relative technology after advancement, so $z_2 < z_1$. One can see from the figure the expansion in the set of potential colonies actually colonized. One can also see that it is reasonable to expect that among later colonies, it is likely that population and inequality are negatively related, while the opposite is true among "earlier" colonies.

We thus have drawn a connection between the initial population density of a (potential) colony and what might be described as the historical level of inequality in that colony. Our model of this process is fairly straightforward and is also capable of explaining how population density and other features of the colonial economy, such as the land intensity of production, interrelate. Unfortunately, historical inequality data are not available to directly evaluate the prediction of a hyperbolic relation between ancient inequality and population density. Instead, following Acemoglu *et al.* (2001), we follow the social conflict view of institutional development, arguing that equilibrium outcomes where a privileged class of landowners is better off than workers provides strong incentives for the élite to defend its position (through a range of political and economic mechanisms). If so, the relation between historical population density and inequality should carry over until today, linking *modern* income distributions to historical population densities in a non-linear fashion. This will be tested in the next section.

4. Empirical Analysis

4.1 Data and descriptive statistics

Our main data requirements regard the measure of inequality – our dependent variable – and historical native population density in former colonies, our main explanatory variable. A frequent point of debate involving past studies on inequality has been the poor quality of the data. Deininger and Squire (1996, 1998) attempted to address these quality issues and proposed a new, expanded dataset, which was updated in 1997. Following Atkinson and Brandolini's (2001) criticism (based on observations for OECD countries from an alternative income survey) of the heterogeneous methodologies used to gather the data presented by Deininger and Squire, the UN's World Institute for Development Economics Research (WIDER) produced a new inequality database, drawing on data from both these previous datasets. The recently updated version (UNU/WIDER 2007) offers some further improvements regarding cross-country comparability and country coverage, including exclusive recent updates to the Deininger and Squire dataset, and more extended information on the survey methodologies and quality and reliability of the data. We report results for one measure of income inequality from this WIDER database, namely the Gini index (as calculated by WIDER according to a new and more accurate method), but all results spill over the analyses using an alternative measure, the top quintile income share, instead (details are available upon request).

<< Table 1 about here >>

For each country, the earliest possible observation – to avoid capturing more recent economic trends (e.g. related to globalization) – of the highest possible quality was chosen that reflected the following fundamental criteria: (i) the income survey covered the entire population; (ii) the basic statistical unit was the household; and (iii) the survey was based on (disposable) income. As our sample of former colonies consists mostly of developing countries, some exceptions had to be made to avoid drastically reducing the number of observations: in one case, the income survey was based on family instead of household units; and for eight countries, the basic statistical unit was listed as uncertain (although the surveys were not generally of bad quality otherwise). Furthermore, for 18 countries there was only consumption or expenditure-based inequality data available.[17]

Data reliability is a further important point. WIDER distinguishes four quality categories according to whether the concepts underlying the survey are known and accepted, and according to the overall quality of the survey. Observations where both quality and concepts are unproblematic receive a rating of "1"; these include entries for virtually all high-income countries, and some more recent survey data for other countries. At the other extreme, observations to be treated as mere "memorandum" entries are classified as quality "4". We were generally able to avoid using worst-quality data on the Gini index, favouring more recent entries for the sake of reliability. The trade-off between data quality and earliest-possible observation for each country resulted in 30% of the Gini index entries stemming from income surveys made after 1980, while the rest come mainly from the 1960s and 1970s, with some

[17] A dummy variable capturing whether the Gini index was based on income or expenditure/consumption data was consistently insignificant in the estimations, showing that results are not driven by different survey bases.

countries offering reliable inequality data for as early as the 1950s.[18] Descriptive statistics for the Gini index are presented in row 1 of Table 1. They show a wide gap between the least and most unequal country—New Zealand having the lowest inequality, and the Central African Republic (CAR) the highest. As a comparison, the world mean Gini index (42.647) is slightly lower than for the colonies sample. However, the most unequal country remains the CAR in the global sample.

Our data on early population densities is taken from McEvedy and Jones (1978) and compiled according to the procedure in Acemoglu et al. (2002). We divide the estimated total population in 1500 by the country's total arable land area in 1995 (from the WDI). Where applicable, we adjust the estimates for the percentage of arable land area (for example in Saharan countries), again using the information contained in McEvedy and Jones (1978). For the most part, our data are analogous to those provided by Acemoglu et al. (2002), with some minor approximation discrepancies, as well as additional observations on non-colonies for our comparison estimations.[19] In both samples, the Seychelles were uninhabited before French colonization, while Egypt had by far the highest population density in 1500 of all former colonies, with over 100 inhabitants per square km on average (row 2 of Table 1).[20] Estimates on population numbers five hundred years ago are close approximations of the truth at best, which likely introduces a bias towards zero into our empirical results. In order to at least

[18] All estimations were performed with the Deininger and Squire (1996, 1998) dataset, as well as period average inequality data, with very similar results, confirming that the observed population density-inequality relationship is not due merely to the chosen dataset (results available upon request).

[19] In a few cases, our data divert from those of Acemoglu et al. (2002). Following the information in McEvedy and Jones (1978), we assume that all Caribbean islands have the same population density of 1.33 (instead of 0.27 for Cuba and 1.46 for others as in Acemoglu et al. (2002)), with the exception of more populous Jamaica, which has a population density of 4.26 in both datasets. We also use the regional population estimates given by McEvedy and Jones (1978) for Malaysia and Singapore to calculate an average population density of 1.21, whereas Acemoglu et al. (2002) assume that Singapore had the same (very low) population density as the United States, namely 0.09. Similarly, our estimate of 9.16 for Hong Kong is based on the notes in McEvedy and Jones (1978), while Acemoglu et al. (2002) again assumed that Hong Kong had the same population density as the United States, even though mainland China in fact had a high average density of 10.72. Robustness tests show that our results are not influenced by the differences.

[20] Egypt was by no means the world's most populous country in 1500: Japan had nearly triple its population density, at an average 287.53 inhabitants per square km.

minimize the influence of outliers while at the same preserving our full sample size, we perform all estimations below using log (population density in 1500 + 1).

We use four different variables to test the "competing" geography hypothesis of the origins of inequality proposed by ES and others, which holds that certain geophysical factors – e.g. a temperate climate, soils suitable to growing staples such as grains, and proximity to the coast – have favored the emergence of more prosperous, egalitarian societies. The first variable which captures one of these geographical aspects, *lwheatsugar,* leans directly on the hypothesis of ES: it measures the soil suitability for growing wheat versus sugarcane, expressed as the (log) ratio of the share of arable land suitable for growing wheat to the share suitable for sugarcane, from Easterly (2007). A lower ratio indicates a propensity to grow a high-yield, plantation-style crop, with the ensuing emergence of a more unequal society, and vice versa. Jamaica shows the greatest soil suitability for growing sugarcane, followed by Guatemala, the Philippines and the Dominican Republic. Next, *latitude* is defined as the absolute distance from the Equator and is taken from La Porta et al. (1999). A greater distance from the Equator is linked to more temperate climates, and to more highly developed countries with less unequal income distributions. A similar concept lies behind the variable *tropics*, which indicates a country's land area in the geographical tropics in percent, while *distcr* measures the mean distance to the nearest coastline or sea-navigable river in km, according to the reasoning that coastal countries tend to be more prosperous due to trade. Both of these last variables are from the Gallup, Mellinger and Sachs geography database (available via the *Center for International Development* website).

As data on early reserves or production of minerals for a large number of countries are virtually impossible to come by, we constructed a dummy variable (*precious*) to capture the possible influence of precious metals on inequality suggested by ES. Countries are assigned value "1" if there was a known presence of gold or silver (or both) in the 16th century, and zero otherwise. Several historical studies served as the basis for this classification, including Del Mar (1902) and Kellenbenz (1981). 29 former colonies from all three regions – the Americas, Africa, and Asia – are reported to have had substantial precious metal deposits at the time.

Finally, we took into account the colonizers' level of (military) technology by constructing a dummy variable for a country colonized with "early" military technology, based on information on the year of colonization from the CIA World Factbook and the U.S. Department of State. We chose a generous cut-off date of 1700 in order to avoid biasing the empirical results in our favour; as described in Section 2, historians in fact most commonly date the period of the "military revolution" between 1560-1660. The 29 countries colonized before the turn of the 18th century are listed as "1", including all of the New World colonies and the Philippines.[21]

4.2 Empirical results

In Table 2, we show some comparison estimations using population density in 1500 to explain modern-day inequality, as measured by the Gini index. For both the entire world and the former colonies sample (columns (1)-(2)), early population density has a strong *negative* influence on inequality, i.e. a higher population density in 1500 is linked to lower inequality today. This finding appears in marked contrast to the reasoning of ES, who argue the exact opposite. However, this simple approach does not adequately capture the whole breadth of the population density-inequality relationship as touched upon in the model and discussion above.

<< Table 2 about here >>

The model suggested opposing effects of population density on settlement decisions of colonists. Our first attempt to capture this is to include a squared term of logged population density (*lpopdens2*). As is evident from columns (3)-(4), this regression supports the inverted-U

[21] As mentioned above, several European countries maintained coastal enclaves in Africa and Asia during the 16th–17th centuries. However, these are not considered as "full" colonies, because the majority of the present-day national land area remained under native control until well into the 18th or even 19th century.

shape predicted by the theoretical model; the pattern becomes particularly clear when we drop Egypt from the sample (column (4)), which is a clear outlier due to its very high early population density However, this analysis suffers from a strong multicollinearity problem between *lpopdens* and *lpopdens2* (correlation 0.94), making the results for the main coefficients of interest unreliable.

Fortunately, there are other ways to capture the colonization process. From Figure 1, it is clear that inequality peaks at mid-range population densities, found in countries colonized around the turn of the 18th century. This period roughly coincides with the end of the military revolution. As we have shown, this enabled Europeans to invade the more populous areas which had before defied conquest. The change in colonization patterns becomes obvious once we separate the colony sample into two subsamples according to the year of colonization. In column (5), we see that the countries colonized with the "early" technology (before 1700) show a strong positive relationship between population density and inequality. This pattern is in accordance with the hypothesis of ES that more populous colonies tended to give rise to extractive-type economic systems with highly unequal societies. It is confirmed empirically for the Americas and the Philippines, and it is carried through to some further colonies in the mid-range of population density as the new technology becomes available.

However, and interestingly, the relationship is *reversed* as more and more populous colonies are added, leading to the strong negative relationship with inequality for the "later" colonies subsample in column (6). The opposite signs and highly significant coefficients in the two subsamples contradict the common wisdom of a linear relationship, instead confirming our own theory of an inverse-U shaped relationship.[22] Up to a certain point, higher native population densities favor forceful control by colonists and the emergence of unequal social and economic structures. But very populous countries will enable less and less aggressive and

[22] The existence of two different subsamples with opposite slopes is confirmed by tests using the clusters approach following Hansen (2000), which moreover does not reject our cut-off date of 1700, albeit preferring the post-industrial 1840 as the cut-off between the subsamples.

all-encompassing colonial policies, resulting in a tendency towards maintaining pre-colonial social structures.

We want to keep the entire sample of former colonies together both because of the relatively small sample size, and in order to test our hypothesis against alternative explanations of the colonial origins of inequality. This can be achieved by adding an interaction term *earlytech*lpopdens* between the technology dummy and early population density, which will assume value 0 for all observations after the military revolution.[23] This presents a simple alternative to the estimations with the quadratic logged term introduced in columns (3)-(4). Our basic regression equation therefore becomes:

$$Inequality = \alpha_0 + \alpha_1 lpopdens + \alpha_2 earlytech * lpopdens + \alpha_3 earlytech + \varepsilon \ . \qquad (37)$$

The coefficient α_2 for *earlytech*lpopdens,* shown in column (6), gives the slope for the countries colonized *before* the military revolution: one can easily see that it is identical to the coefficient in column (2) by subtracting the coefficient α_1 for the later-colonized countries, 13.544 – 3.672 = 9.872.

Together, our variables can account for one fourth of the variation in inequality in all former colonies, which is considerable in light of the complexity of the issue. The results also suggest that up to the turning point, increasing (log) population density by one standard deviation (0.93) produces a corresponding one-standard-deviation increase in inequality ((0.93*9.872)/9.168). After the "cusp" however, the same change in (log) population density *decreases* inequality by 0.37 of a standard deviation ((0.93*-3.672)/9.168).

<< Table 3 about here >>

Using this strategy, we now test our hypothesis against other variables. In Table 3, we one-by-one add variables corresponding to the other two main explanations for inequality

[23] This simple method of taking into account differing slope coefficients in subsamples by adding an interaction with a dummy variable—sometimes termed a "structural break" model—was first formally described in Suits (1957). See e.g. Kennedy (2003: Ch. 14) for further details on this technique.

proposed by ES, namely geography and the exploitation of precious metals. The "horse race" of the explanations on the colonial origins of inequality clearly shows that our hypothesis holds up very well to the test. The early population density coefficients remain consistently highly significant and with opposite signs, with little variation in the magnitude of the impact on inequality, even when we add all variables together in column (6).

As far as the other variables are concerned, the wheat-sugar ratio and distance to the sea or a navigable river are not good predictors of inequality, although they each enter with the expected sign. Tropical location is weakly negatively significant only when we simultaneously control for all other effects. The only geographic measure which holds its own against the early population density variables is latitude, suggesting that greater proximity to the Equator does have an independent negative impact on egalitarian income distribution in former colonies.

Interestingly, the early presence of precious metals has a relatively strong effect on inequality, apparently confirming the hypothesis of ES that mineral exploitation tended to produce more unequal societies. However, this effect is not robust to controlling for geographical factors. In separate estimations (available upon request), we also controlled for further effects, including regions, colonial and legal origin, and ethnic fractionalization, none of which influenced our results.

<< Table 4 about here >>

It is possible that our results are being influenced by the presence of outliers. The most commonly cited outliers in the literature on colonial roots of economic development are the so-called "neo-Europes" – Canada, the U.S., Australia and New Zealand. In our case, one could well imagine that their inclusion would bias the results in our favor, as they were all relatively sparsely populated around 1500 and demonstrate low modern inequality levels. Tests confirm this possibility: New Zealand, Canada and Australia in particular often appear as outliers.[24] In Table 4, we therefore repeat our main estimations without these four countries. All of our previous findings regarding the inverse-U shaped relationship between early population density and inequality are confirmed, while the influence of the geographic variables and the precious metals dummy is further weakened. Moreover, the regression fit improves to 0.32 in our basic estimation (column 1). Finally, and as mentioned above, all these results spill over to regression specifications using the income share of the richest quintile of the population to measure inequality.

4.3 Linking inequality to growth and income levels

Finally, we briefly test the usefulness of the analysis by revisiting the contested relationship between inequality on the one hand, and growth and development on the other. As argued above, the theoretical and empirical literature on the nature of this relation is divided. Our analysis suggests exploring the potential of using historical population density, combined with the interaction term, as instruments to predict inequality in the former colonies. To illustrate the importance of careful instrument selection, we also demonstrate similar results for a naïve model based on a simple linear relation between historic population density and inequality. Representative results for the second-stage coefficients for inequality measured by the Gini index are reported in Table 5.

<< Table 5 about here >>

Two results stand out. First, when using our preferred set of instruments, there does not appear to be a causal link from inequality to growth in the former colonies. There is a significant positive link between inequality and current income levels (column 1), but this disappears once we control for regional effects (not shown). Perhaps the various positive and

[24] Further outliers indicated by the tests (using studentized residuals and overall measures of influence) include Laos, Jamaica, Brazil, and Suriname. All of our results are reinforced when dropping these countries one at a time, as well as collectively.

negative effects tend to offset and neutralize each other, the implication being that any correlation between inequality and growth (or income levels) is driven by either reverse causality or omitted variables. This result contrasts the findings of Easterly (2007), who found that inequality was detrimental for income. We believe this shows that what holds for a worldwide sample (as in Easterly's case) need not necessarily be true in a subsample – even a large and important one such as former colonies.25

Second, the choice of instruments matters for the results of the second stage. The naïve model, based on a linear relation in the first stage, delivers weaker results, and the first stage test statistics are generally weaker and show that the explanatory power of the more complex model with the interaction term is superior.

5. Conclusions

The distribution of income is a key topic in economic research, and an important issue for policymakers. It is of interest in and of itself, but it is also relevant because inequality has been linked to enhanced or depressed economic growth rates. This paper examines the historical origins of modern income distribution, and revisits the link between inequality and income (and growth).

Our story is related to that of economic historians Engerman and Sokoloff (1997, 2000, 2002), who focused on the importance of colonial-era factor endowments for development paths in the Americas. They argued that a climate and soil conducive to growing high-value plantation crops such as sugarcane; the existence of gold and silver mines; and a large native population providing cheap labor tended to bring forth social and political structures which favored the colonial élites. The institutional set-ups created to guarantee the status of the

[25] In fact, applying the naïve model with early population density as the sole instrument for inequality to the global sample (129 countries) confirms Easterly's result of a significant negative relationship between inequality and income (coefficient of -0.036, p-value of 0.026).

privileged classes influenced economic development and led to the persistence of these early forms of inequality well into the modern period.

However, we argue that the crucial influence of native population density on European colonization and the subsequent distribution of wealth was not uniform across all colonies. Instead, it was determined by the ratio of land to native inhabitants which the early colonizers were confronted with, and by the (military) technology of which they disposed. Colonizers traded off a higher native population, with its promise of cheap labor and high land rents, against the cost of military conflict and control. Using a model where (potential) colonists can choose between peaceful settlement or aggressive conquest, we demonstrate the existence of an inverse-U shaped relationship between population density at the time of colonization and the income distribution in former colonies. The relative technological level of the colonists plays an important role, as a better technology permits the forceful occupation and subjugation of the more densely populated areas, with the possibility of earning higher land rents.

The findings of the model are tested empirically for a broad sample of former colonies by regressing native population densities around 1500 on modern-day income distributions. The empirical results confirm the inverse-U shaped link between early population density and inequality, as well as the role of the interaction between population density in the colonies and the military technology level of the colonizers in determining subsequent development patterns. We show that early native population density was more important in determining European colonization than geographical factors (soil quality, climate, access to a sea-navigable port, or the presence of precious metals), and use the results to motivate the choice of novel instruments for inequality in regressions explaining income growth or levels. Our 2SLS regressions suggest that there is no strong causal relation running from inequality to growth or income levels; if anything, higher inequality seems to be linked to higher present-day income levels in former colonies.

The results suggest that no one theory of the influence of colonial factor endowments on development patterns can fit all countries, but that there are some important non-linearities which must be considered when trying to understand the origins of economic development.

Appendix 1

A first step in solving the model is to develop expressions for aggregate income used in the demand functions in (3). Since all production functions are homogenous of degree one, incomes can be written as a simple sum of factor payments. Using *w* to denote wages and *r* to denote the rental rates on land, in the potential colony, we have:

$$I_n = w_h L_{nh} + w_y L_{ny} + r_h A_{nh} + r_y A_{ny} \tag{A.1}$$

While in the colonizing country, total income is:

$$I_c = w_c L_c + r_y A_{cy} + w_y L_{cy}\,. \tag{A.2}$$

Plugging these expressions into the demands in (3) results in the following expressions for the demands for goods *Y* and *H*; the model can be solved in terms of these goods, and Walras's law can be relied upon the determine the price of good *X*.

$$Y^d = \alpha \frac{(w_x L_x + r_y A_{cy} + w_y L_{cy})}{p_y}. \tag{A.3}$$

$$H^d = \frac{\delta(w_h L_{nh} + w_y L_{ny} + r_h A_{nh} + r_y A_{ny})}{p_h} \tag{A.4}$$

Because in equilibrium wages and rental rates are given by marginal value products, we can use expressions for marginal value products to eliminate wages and rental rates in (A.3) and (A.4). Using the home production technology (7), we have the wage rate in the colonial power as:

$$w_x = p_x \theta_x\,, \tag{A.5}$$

In the colony, we have the following expressions for wages and rental rates:

$$w_y = p_y(1-\beta_y)\frac{Y}{L_y},\quad r_y = p_y\beta_y\frac{Y}{A_y}$$

$$w_h = p_h(1-\beta_h)\frac{H}{L_h},\quad r_h = p_h\beta_h\frac{H}{A_h}. \tag{A.6}$$

Combining marginal value products in (A.6) with demand in (A.3) gives the following expression for the total demand for good *Y*:

$$Y\left(1-\alpha\beta_y\frac{A_{cy}}{A_y}-\alpha(1-\beta_y)\frac{L_{cy}}{L_y}\right)=\frac{\alpha p_x X}{p_y}, \tag{A.7}$$

The same methodology gives us the demand for good *H* as:

$$H = \frac{(1-\delta)(p_h H + p_y(1-\beta_y)\frac{Y}{L_y}L_{ny} + p_y\beta_y\frac{Y}{A_y}A_{ny})}{p_h}\,. \tag{A.8}$$

Solving (A.7) for the price of Y and using the normalization that $p_x = 1$, we have the equilibrium real price of Y as:

$$p_y^* = \frac{p_y}{p_x} = \frac{X}{Y}\frac{\alpha}{\left(1-\alpha\beta_y\frac{A_{cy}}{A_y}-\alpha(1-\beta_y)\frac{L_{cy}}{L_y}\right)} \tag{A.9}$$

Solving (A.8) for the price of H in terms of Y yields:

$$\frac{p_h}{p_y} = \frac{Y}{H}\frac{(1-\delta)}{\delta}\left((1-\beta_y)\frac{L_{cy}}{L_y}+\beta_y\frac{A_{cy}}{A_y}\right) \tag{A.10}$$

Using (A.9) and (A.10), we can write the price of good H in terms of good X as:

$$p_h^* = \frac{X}{H}\frac{\alpha\delta\left(\beta_y\frac{A_{ny}}{A_y}+(1-\beta_y)\frac{L_{ny}}{L_y}\right)}{(1-\delta)\left(1-\alpha\beta_y\frac{A_{cy}}{A_y}-\alpha(1-\beta_y)\frac{L_{cy}}{L_y}\right)}. \tag{A.11}$$

The equilibrium conditions described in (10) can be written:

$$p_y(1-\beta_y)\frac{Y}{L_y} = p_h(1-\beta_h)\frac{H}{L_h}; \qquad p_y\beta_y\frac{Y}{A_y} = p_h\beta_h\frac{H}{A_h}. \tag{A.12}$$

The equilibrium conditions described in (12) and (13) can be written as:

$$p_y(1-\beta_y)\frac{Y}{L_y} = \theta_x; \quad p_y\beta_y\frac{Y}{A_y}\frac{A}{D+M} = \theta_x. \tag{A.13}$$

Plugging the prices from (A.9) and (A.11) into (A.12) and (A.13), results in the following system of equations:

$$\frac{\delta}{1-\delta}\left(\beta_y\frac{A_{ny}}{A_y}+(1-\beta_y)\frac{L_{ny}}{L_y}\right)\frac{(1-\beta_h)}{L_h} = \frac{(1-\beta_y)}{L_y}; \tag{A.14}$$

$$\frac{\delta}{1-\delta}\left(\beta_y\frac{A_{ny}}{A_y}+(1-\beta_y)\frac{L_{ny}}{L_y}\right)\frac{\beta_h}{A_h} = \frac{\beta_y}{A_y} \tag{A.15}$$

$$\frac{\alpha X}{1-\alpha\beta_y\frac{A_{cy}}{A_y}-\alpha(1-\beta_y)\frac{L_{cy}}{L_y}}\frac{1-\beta_y}{L_y} = \theta_x \tag{A.16}$$

$$\frac{\alpha X}{1-\alpha\beta_y\frac{A_{cy}}{A_y}-\alpha(1-\beta_y)\frac{L_{cy}}{L_y}}\frac{\beta_y}{A_y}\frac{A}{D+M} = \theta_x \tag{A.17}$$

The conditions (A.14) and (A.15) represent the allocation problem of the native population in the colony, while (A.16) and (A.17) represent the allocation problem of the colonizing power. Solving (A.14-17) requires endowment constraints. Endowment constraints are:

$$L_n = L_{ny} + L_{nh}; \qquad A_n = A_{ny} + A_{nh};$$

$$L_c = L_{cy} + M + L_x; \qquad A_n = A - A_{cy}. \tag{A.18}$$

Furthermore:

$$A_y = A_{cy} + A_{ny}; \qquad L_y = L_{cy} + L_{ny}. \tag{A.19}$$

The expressions A_{cy} and A_n are described in the text (equations 8 and 9) as:

$$A_{cy} = \frac{AM}{M+D}; \qquad A_n = \frac{AD}{M+D}. \tag{A.20}$$

Using (A.18), total output of good X can be written as $X = \theta_x(L_c - M - L_{cy})$. Using (A.18), (A.19), and (A.20) in (A.14-16) results in a four equation system with four unknowns: $M, L_{ny}, L_{cy}, A_{ny}$. Solving the system for these four unknowns gives the solutions shown in the text (15-18).

Appendix 2

Given the utility function in (39), one finds that demand functions are now given by:

$$Y(i) = \frac{\alpha I_c}{p_y(i)}, \tag{A.21}$$

$$H(i) = \frac{\delta I_n(i)}{p_h(i)},\ i \in [0,1]. \tag{A.22}$$

As before, we rely on Walras's law to determine the demand for the good *X*. Given some subset of colonies $i \le i^*$ are actually colonized, total income for the colonial power may be written as:

$$I_c = w_x L_x + \int_0^{i^*} w_y(i) L_{cy}(i) di + \int_0^{i^*} r_y(i) A_{cy}(i) di. \tag{A.23}$$

Whereas income in each particular colony can be denoted as:

$$I_n(i) = w_h(i) L_h(i) + w_y(i) L_{ny}(i) + r_h(i) A_{nh}(i) + r_y(i) A_{ny}(i). \tag{A.24}$$

Plugging income (A.23) into the demand function for good Y (A.21) results in:

$$p_y(i) y(i) = \alpha\left(w_x L_x + \int_0^{i^*} w_y(i) L_{cy}(i) di + \int_0^{i^*} r_y(i) A_{cy}(i) di \right) \tag{A.25}$$

A consequence of (A.25) is that in equilibrium $p_y(i) y(i) = p_y(j) y(j)$. Total expenditures on any two colonial goods are equal.

Wages and rental rates must be equal to marginal value of product in equilibrium. This requires that (taking into account that we have assumed $\beta_h = \beta_y = \frac{1}{2}$):

$$w_h(i) = \frac{1}{2} p_h(i) \frac{H(i)}{L_h(i)}, \qquad r_h(i) = \frac{1}{2} p_h(i) \frac{H(i)}{A_h(i)},$$

$$w_y(i) = \frac{1}{2} p_y(i) \frac{Y(i)}{L_y(i)}, \qquad r_y(i) = \frac{1}{2} p_y(i) \frac{Y(i)}{A_y(i)}, \tag{A.26}$$

Using (A.26) in (A.25), and applying the result that $p_y(i) y(i) = p_y(j) y(j)$, we arrive at:

$$p_y(i) = \frac{w_x L_x}{Y(i)} \frac{\alpha}{1 - \frac{\alpha}{2} \int_0^{i^*} \left(\frac{L_{cy}(j)}{L_y(j)} + \frac{A_{cy}(j)}{A_y(j)} \right) dj}. \tag{A.27}$$

Since $w_x L_x = p_x X$, and $p_x = 1$ by assumption, the demand functions in (A.27) become:

$$p_y(i) = \frac{X}{Y(i)} \frac{\alpha}{1 - \frac{\alpha}{2} \int_0^{i^*} \left(\frac{L_{cy}(j)}{L_y(j)} + \frac{A_{cy}(j)}{A_y(j)} \right) dj}. \tag{A.28}$$

In any particular colony, we have, after substituting the expression for income (A.24) into the demand function (A.22):

$$H(i)p_h(i) = \delta\big(w_h(i)L_h(i) + w_y(i)L_{ny}(i) + r_h(i)A_{nh}(i) + r_y(i)A_{ny}(i)\big). \qquad \text{(A.29)}$$

Using our expressions for wages and rental rates in equilibrium (A.26), we arrive at:

$$H(i) = \frac{\delta\left(p_h(i)H(i) + p_y(i)\frac{1}{2}\frac{Y(i)}{L_y(i)}L_{ny}(i) + p_y(i)\frac{1}{2}\frac{Y(i)}{A_y(i)}A_{ny}(i)\right)}{p_h(i)}. \qquad \text{(A.30)}$$

Simplifying and rearranging (A.30), we find:

$$\frac{p_h(i)}{p_y(i)} = \frac{1}{2}\frac{Y(i)}{H(i)}\frac{\delta}{1-\delta}\left(\frac{L_{ny}(i)}{L_y(i)} + \frac{A_{ny}(i)}{A_y(i)}\right). \qquad \text{(A.31)}$$

Using our expression for the price of good *Y* in (A.28), (A.31) becomes:

$$p_h(i) = \frac{1}{2}\frac{X}{H(i)}\frac{\delta\alpha}{1-\delta}\frac{\left(\frac{L_{ny}(i)}{L_y(i)} + \frac{A_{ny}(i)}{A_y(i)}\right)}{1-\frac{\alpha}{2}\int_0^N\left(\frac{L_{cy}(j)}{L_y(j)} + \frac{A_{cy}(j)}{A_y(j)}\right)dj} \qquad \text{(A.32)}$$

Within each colony, we require that marginal value products for land and labor are equal across production of goods *H* and *Y*. That is:

$$p_y(i)\frac{Y(i)}{A_y(i)} = p_h(i)\frac{H(i)}{A_h(i)}, \qquad \text{(A.33)}$$

$$p_y(i)\frac{Y(i)}{L_y(i)} = p_h(i)\frac{H(i)}{L_h(i)} \qquad \text{(A.34)}$$

Using (A.33) and (A.34), along with the expressions for equilibrium prices (A.28) and (A.32) results in the following equations describing factor endowments within an individual colony:

$$\frac{\delta}{1-\delta}\frac{1}{2}\left(\frac{L_{ny}(i)}{L_y(i)} + \frac{A_{ny}(i)}{A_y(i)}\right)\frac{1}{A_h(i)} = \frac{1}{A_y(i)}, \qquad \text{(A.35)}$$

$$\frac{A_y(i)}{L_y(i)} = \frac{A_h(i)}{L_h(i)}. \qquad \text{(A.36)}$$

Coupled with these two conditions, we have the constraints:

$$A_h(i) = A_n(i) - A_{ny}(i), \qquad \text{(A.37)}$$

$$A_n(i) = A(i) - A_{cy}(i), \qquad \text{(A.38)}$$

$$L_h(i) = L_n(i) - L_{ny}(i), \qquad \text{(A.39)}$$

$$L_y(i) = L_{ny}(i) + L_{cy}(i), \tag{A.40}$$

$$A_y(i) = A_{cy}(i) + A_{ny}(i), \tag{A.41}$$

Solving for equilibrium allocations, taking as given colonial decisions, results in the following native resources allocated to production of good *Y*:

$$A_{ny}(i) = [A(i) - A_{cy}(i)](1-\delta) - \frac{\delta A(i)}{2}\frac{L_n(i)}{L_n(i)+L_{cy}(i)}; \tag{A.42}$$

$$L_{ny}(i) = L_n(i)(1-\delta) - \frac{\delta L_{cy}(i)}{2} + \frac{\delta(L_n(i)+L_{cy}(i))}{2}\frac{A_{cy}(i)}{A(i)} \tag{A.43}$$

Total factors allocated to good Y then become:

$$A_y(i) = (1-\delta)A(i) + \frac{1}{2}\delta A_{cy}(i) + \frac{\delta A(i)}{2}\frac{L_{cy}(i)}{L_n(i)+L_{cy}(i)} \tag{A.44}$$

$$L_y(i) = L_n(i)(1-\delta) + \left(1-\frac{\delta}{2}\right)L_{cy}(i) + \frac{\delta(L_n(i)+L_{cy}(i))}{2}\frac{A_{cy}(i)}{A(i)}. \tag{A.45}$$

The colonial land acquisition function is given by:

$$A_{cy}(i) = \frac{A(i)M(i)}{M(i)+D(i)}, \tag{A.46}$$

Inserting this into each of the above conditions results in expressions (A.42-46) as:

$$A_{ny}(i) = A(i)\left(\frac{D(i)}{M(i)+D(i)} - \frac{\delta}{2}\frac{(2L_n(i)D(i)+D(i)L_{cy}(i)+M(i)L_n(i)}{(L_n(i)+L_{cy}(i))(D(i)+M(i))}\right); \tag{A.47}$$

$$L_{ny}(i) = L_n(i)(1-\delta) - \frac{\delta}{2}L_{cy}(i) + \frac{\delta}{2}\frac{M(i)}{D(i)+M(i)}(L_n(i)+L_{cy}(i)) \tag{A.48}$$

$$A_y(i) = A(i)\left(1-\frac{\delta}{2}\frac{2L_n(i)D(i)+D(i)L_{cy}(i)+M(i)L_n(i)}{(D(i)+M(i))(L_n(i)+L_{cy}(i))}\right) \tag{A.49}$$

$$L_y(i) = L_{cy}(i) + L_n(i) - \frac{\delta}{2}\frac{M(i)L_n(i)+D(i)L_{cy}(i)+2L_n(i)D(i)}{D(i)+M(i)} \tag{A.50}$$

This completes the description of how resources are allocated within a colony, given some level of colonial intervention.

We now turn to invoking our equilibrium conditions describing colonial behavior. Within each country, we require that the marginal value product from colonial effort in productive labor is equal to the marginal value product of effort in seizing land. Suppose that some aggregate amount of labor $C(i)$ has been allocated to colonizing country i. Our equilibrium condition is that the marginal product of labor across the two activities should be equal:

$$w_y(i) = r_y(i)a(i). \tag{A.51}$$

Condition (A.51) becomes, after using expressions for equilibrium marginal value products in (A.26),

$$\frac{1}{L_y(i)} = \frac{a(i)}{A_y(i)}. \tag{A.52}$$

Since $a(i) = \frac{A(i)}{D(i) + M(i)}$, (A.52) becomes:

$$\frac{1}{L_y(i)} = \frac{1}{A_y(i)} \frac{A(i)}{D(i) + M(i)}. \tag{A.53}$$

Inserting $A_y(i)$ from (A.49) and $L_y(i)$ from (A.50) into (A.53) and solving, we arrive at optimal colonial interventions as:

$$L_{cy}(i) = \tfrac{1}{2}(C(i) + D(i) - L_n(i)), \tag{A.54}$$

$$M(i) = \tfrac{1}{2}(C(i) - D(i) + L_n(i)). \tag{A.55}$$

We now bring in the assumption that $D(i) = L_n(i) = \frac{1}{2} zN(i)$. Under this assumption, we find that the eventual share of land in production of Y can be written as:

$$\frac{A_{cy}(i)}{A_y(i)} = \frac{C(i)}{(1-\delta)zN(i) + C(i)}. \tag{A.56}$$

While the equilibrium share of labor is given by:

$$\frac{L_{cy}(i)}{L_y(i)} = \frac{C(i)}{(1-\delta)zN(i) + C(i)}. \tag{A.57}$$

Having determined equilibrium shares given allocations of effort within countries, we now determine the pattern of colonial effort across the colonies. This requires that wages and rental rates earned by colonists be equal in every potential colony. Using expressions for marginal value products (A.26), equilibrium prices (A.28) and (A.32), and the solutions (A.54) and (A.55), we find that:

$$w_y(i) = \frac{\alpha X}{1 - \alpha \int_0^{i^*} \frac{C(j)}{(1-\delta)zN(j) + C(j)} dj} \frac{1}{(1-\delta)zN(i) + C(i)} \tag{A.58}$$

In equilibrium, we require that, for any two colonies, $w_y(i) = w_y(j)$. From (A.58), this requires that

$$(1-\delta)zN(j)+C(j)=(1-\delta)zN(i)+C(i). \quad \text{(A.59)}$$

Consider now the "cutoff" potential colony. This is the first potential colony for which there is no colonization; that is, $C(i^*)=0$. Substituting this into (A.59) results in:

$$C(j)=[N(i^*)-N(j)](1-\delta)z. \quad \text{(A.60)}$$

Total production of good X is given by $X=\theta_x\left(L_c-\int_0^{i^*}C(j)dj\right)$. Forming the expression $w_x=w_y(i)$ using (A.58) gives:

$$\frac{\alpha\left(\theta_x(L_c-\int_0^{i^*}C(i)di\right)}{(1-\delta)zN(i^*)}=\theta_x\left(1-\alpha\int_0^{i^*}\frac{C(j)}{(1-\delta)zN(j)+C(j)}dj\right). \quad \text{(A.61)}$$

Applying (A.60) in the denominator of the term inside the integral on the right-hand side of (A.61) results in:

$$\frac{\alpha\left(L_c-\int_0^{i^*}C(i)di\right)}{(1-\delta)zN(i^*)}=\left(1-\alpha\int_0^{i^*}\frac{C(j)}{(1-\delta)zN(i^*)}dj\right). \quad \text{(A.62)}$$

Solving (A.62) results in

$$N(i^*)=\frac{\alpha L_c}{(1-\delta)z}. \quad \text{(A.63)}$$

Since $N(i)=\eta i$, we have:

$$i^*=\frac{\alpha L_c}{(1-\delta)z\eta} \quad \text{(A.64)}$$

Using (A.64) in (A.60) results in:

$$C(j)=\eta[i^*-j](1-\delta)z. \quad \text{(A.65)}$$

Given the solution in (A.65), we see that eventual factor shares of the colonists in each country colonized become as given in the text, equation (41). Total income in each country is

$$I_n(i)=p_h(i)H(i)+p_y(i)Y(i), \quad \text{(A.66)}$$

The total amount of income held by natives is given by:

$$I_n(i)=p_h(i)H(i)+p_y(i)Y(i)\frac{L_{ny}(i)}{L_y(i)}. \quad \text{(A.68)}$$

Therefore, our expression for the income side of the gini coefficient becomes:

$$s_I(i) = \frac{p_h(i)H(i) + p_y(i)Y(i)\dfrac{L_{ny}(i)}{L_y(i)}}{p_h(i)H(i) + p_y(i)Y(i)}. \tag{A.69}$$

Plugging in equilibrium prices (A.28) and (A.32), we have:

$$s_I(i) = \frac{L_{ny}(i)}{\delta L_{ny}(i) + (1-\delta)L_y(i)}, \tag{A.70}$$

Using equilibrium factor shares, (A.70) becomes:

$$s_I(i) = \frac{i}{\delta i + (1-\delta)i^*}. \tag{A.71}$$

Given that total intervention in the *i*th colony in (A.65), we find that the eventual native population share is

$$s_p(i) = \frac{\eta i}{\eta[i^* - j](1-\delta)z + \eta i} = \frac{i}{[i^* - i](1-\delta)z + i}. \tag{A.72}$$

References

Acemoglu, D., S. Johnson, and J. A. Robinson, 2001. The colonial origins of comparative development: an empirical investigation. *American Economic Review* 91: 1369-1401.

Acemoglu, D., S. Johnson, and J. A. Robinson, 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics* 117: 1231-1294.

Acemoglu, D. and J.A. Robinson, 2006. De facto political power and institutional persistence. *American Economic Review* (Papers and Proceedings) 96: 325-330

Alesina, A. and D. Rodrik, 1994. Distributive politics and economic growth. *Quarterly Journal of Economics* 108: 465-490

Anand, S. and S. Kanbur, 1993a. The Kuznets process and the inequality-development relationship. *Journal of Development Economics* 40: 25-52

Anand, S. and S. Kanbur, 1993b. Inequality and development: A critique. *Journal of Development Economics* 41: 19-43

Atkinson, A.B., and A. Brandolini, 2001. Promises and pitfalls in the use of secondary data-sets: income inequality in OECD countries as a case study. *Journal of Economic Literature* 39 (3): 771-800.

Baker, M.J., 2003. An equilibrium model of land tenure in hunter-gatherer societies. *Journal of Political Economy* 111: 124-173.

Black, J., 1995. A Military Revolution? In: Rogers, C. J. (ed). *The Military Revolution Debate: Readings on the Military Transformation of Early Modern Europe*. Boulder: Westview Press.

Bolton, H. E., and T. M. Marshall, 1920. *The colonization of North America 1492-1783*. New York: Macmillan.

Bourguignon, F. and C. Morrisson, 1998. Inequality and development: The role of dualism. *Journal of Development Economics* 57: 233-257

Bourguignon, F. and T. Verdier, 2000. Oligarchy, democracy, inequality and growth. *Journal of Development Economics* 62: 285-313

Cable, J., 1998. *The Political Influence of Naval Force in History*. London: Macmillan Press

Clarke, G., 1995. More evidence on income distribution and growth. *Journal of Development Economics* 47: 403-428

Coatsworth, J. H., 1999. Economic and institutional trajectories in 19th century Latin America. In: J.H. Coatsworth and A.M. Taylor (eds.): *Latin America and the world economy since 1800.*Cambridge, MA: Harvard University Press.

Del Mar, A., 1902. *A History of the Precious Metals from the Earliest Times to the Present*. New York: Burt Franklin, 2nd edition.

Deininger, K. and L. Squire, 1996. A new data set measuring income inequality. *World Bank Economic Review* 10 (3), 565-591.

Deininger, K. and L. Squire, 1998. New ways of looking at old issues: inequality and growth. *Journal of Development Economics* 57 (2), 259-287.

Easterly, W., 2007. Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics* 84 (2): 755-776.

Easterly, W., and R. Levine, 2003. Tropics, germs, and crops: how endowments influence economic development. *Journal of Monetary Economics* 50: 3-39.

Engermann, S. L., and K. L. Sokoloff, 1997. Factor endowments, institutions, and differential paths of growth among New World economies. In: S.H. Haber (ed.). *How Latin America fell behind*. Stanford, CA: Stanford University Press.

Engermann, S. L., and K. L. Sokoloff, 2002. Factor Endowments, Inequality, and Paths of Development among New World Economies. NBER Working Paper No. 9259.

Ferguson, N., 2002. *Empire: the rise and demise of the British world order and the lessons for global power*. London: Allen Lane.
Ferro, M. 1997. *Colonization: a global history*. London: Routledge.
Forbes, K., 2000. A reassessment of the relationship between inequality and growth. *American Economic Review* 90: 869-887
Förster, S., W. J. Mommsen, and R. Robinson (eds.), (1988). *Bismarck, Europe, and Africa: the Berlin Africa Conference 1884-1885 and the onset of partition*. Oxford: Oxford University Press.
Gellar, S., 1995. The colonial era. In: P.M. Martin and P. O'Meara (eds.), *Africa*. Bloomington, IN: Indiana University Press, 3rd edition.
Glaeser, E.L., R. La Porta, F. Lopez-De-Silanes and A. Shleifer, 2004. Do Institutions Cause Growth? *Journal of Economic Growth* 9: 271-303
Glete, J. (ed.), 2005. *Naval history 1500-1680*. Aldershot: Ashgate.
Grossman, H., and M. Kim, 1995. Swords or Plowshares: A Theory of the Security to Claims of Property. *Journal of Political Economy* 103: 1275-1288
Guilmartin, J.F., Jr., 1995. The Military Revolution: Origins and First Tests Abroad. In: Rogers, C. J. (ed). *The Military Revolution Debate: Readings on the Military Transformation of Early Modern Europe*. Boulder: Westview Press.
Hansen, B.E., 2000. Sample splitting and threshold estimation. *Econometrica* 68: 575-603.
Hirschleifer, J. 1995. Anarchy and its Breakdown. *Journal of Political Economy* 103, pp. 26-52
Kaldor, N., 1956. Alternative theories of distribution. *Review of Economic Studies* 23: 94-100
Kellenbenz, H. (ed.), 1981. *Precious Metals in the Age of Expansion*. Stuttgart: Klett-Cotta.
Keim, C. A., 1995. Africa and Europe before 1900. In: P.M. Martin and P. O'Meara (eds.), *Africa*. Bloomington, IN: Indiana University Press, 3rd edition.
Kennedy, P., 2003. *A guide to econometrics*. Oxford: Blackwell Publishing, 5th edition.
Kuznets, S., 1955. Economic growth and income inequality. *American Economic Review* 45: 1-28
La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. W. Vishny, 1999. The quality of government. *Journl of Law, Economics, and Organization* 15(1): 222-279.
Lewis, W., 1954. Economic development with unlimited supplies of labor. *The Manchester School of Economics and Social Studies* XXII: 139-191.
Louis, W. R. (ed.), 1976. *Imperialism: the Robinson and Gallagher controversy*. New York: Franklin Watts.
Lynn, J. A., 1991. Clio in arms: the role of the military variable in shaping history. *Journal of Military History* 55 (1): 83-95.
Mariscal, E. and K. L. Sokoloff, 2000. Schooling, Suffrage, and the Persistence of Inequality in the Americas, 1800-1945. In: S. Haber. (ed.), *Political Institutions and Economic growth in Latin America: Essays in Policy, History, and Political Economy*. Stanford, CA: Hoover Institution Press.
McEvedy, C. and R. Jones, 1978. *Atlas of world population history*. London: Lane.
North, D. C., W. Summerhill, and B. Weingast, 2000. Order, Disorder and economic change: Latin America vs. North America. In: B. Bueno de Mesquita and H. L. Root (eds.), *Governing for Prosperity*. New Haven, CT: Yale University Press.
Parker, G., 1996. *The Military Revolution: Military innovation and the rise of the West 1500-1800*. New York: Cambridge University Press, 2nd edition.
Persson, T. and G. Tabellini, 1994. Is inequality harmful for growth? *American Economic Review* 84: 600-621
Roberts, M., 1995. The Military revolution, 1560-1660. In: C.J. Rogers, Jr. (ed.). *The military revolution debate: readings on the military transformation of early modern Europe*. Boulder: Westview Press.

Robinson, R. (1972). Non-European foundations of European imperialism: sketch for a theory of collaboration. In: E.R.J. Owen and R. B. Sutcliffe, (eds.). *Studies in the theory of imperialism*. London: Longman.

Rogers, C. J., Jr. (ed.), 1995. *The military revolution debate: readings on the military transformation of early modern Europe.* Boulder: Westview Press.

Sokoloff, K. L., and S. L. Engermann, 2000. Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives* 14 (3): 217-232.

Suits, D. B., 1957. Use of dummy variables in regression equations. *Journal of the American Statistical Association* 52 (280), 548-551.

UNU/WIDER, 2007. World Income Inequality Database. Version 2.0b, May 2007.

Waites, B., 1999. *Europe and the Third World: From Colonisation to Decolonisation c. 1500-1998.* London: Macmillan.

Wittfogel, K. A., 1957. *Oriental despotism: A comparative study of total power*. New Haven, CT: Yale University Press.

Young, C., 1994. *The African colonial state in comparative perspective*. New Haven, CT: Yale University Press.

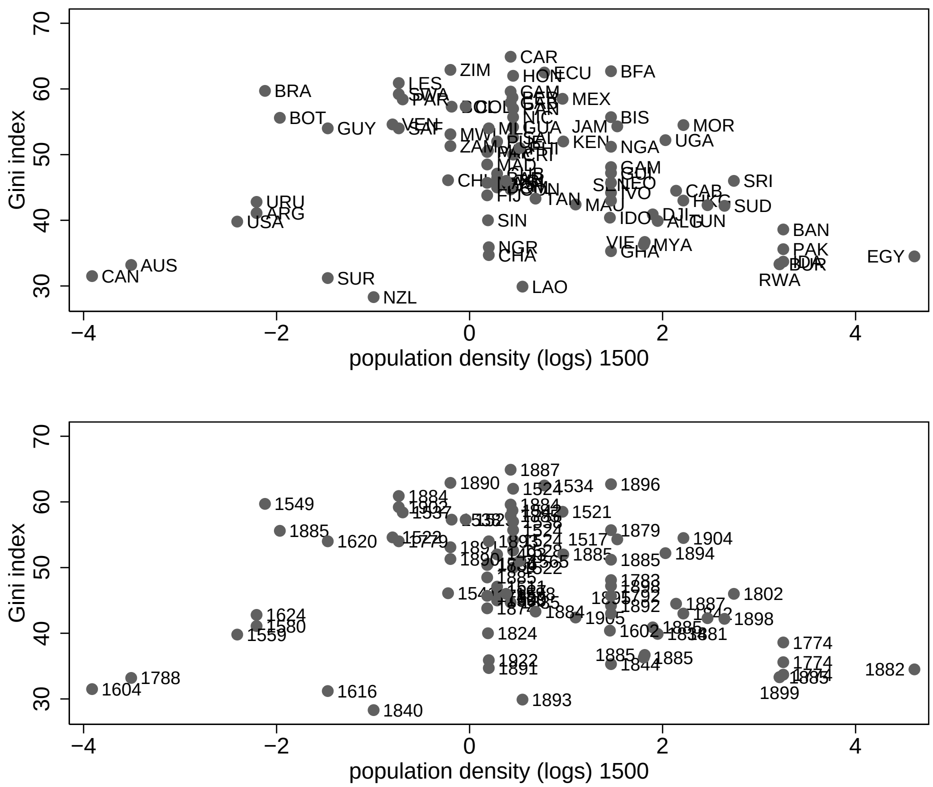

Figure 1: Scatter plot of (log) population density in 1500 and modern inequality

Notes: Inequality is measured by the Gini index. Figure 1(a) above shows results listed by colony, Figure 1(b) below by year of colonization.

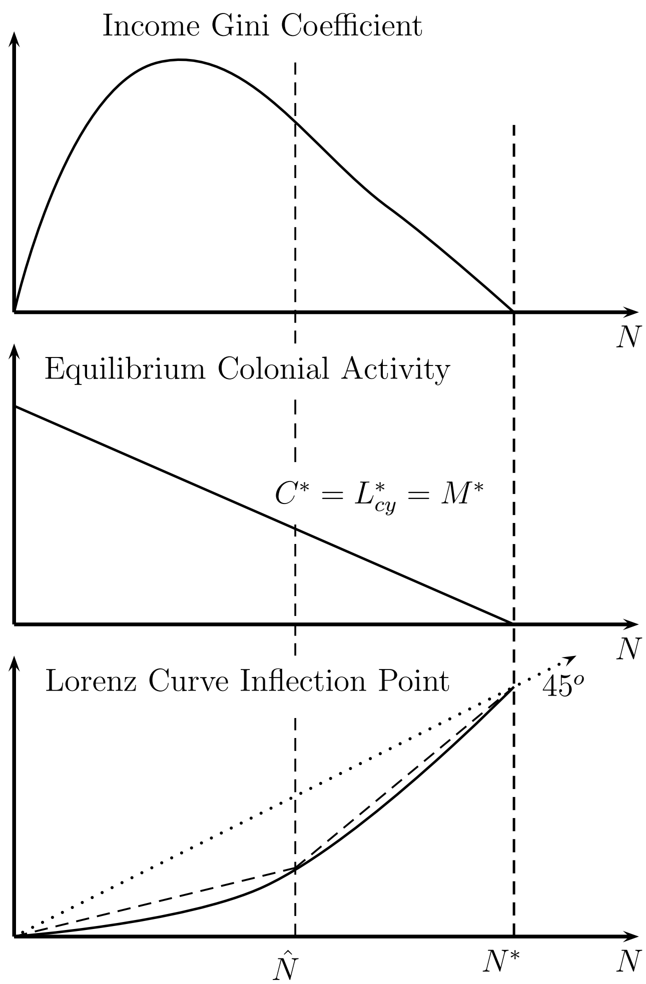

Figure 2: The income Gini coefficient, equilibrium settlement behavior, and the Lorenz curve inflection point as functions of initial population density

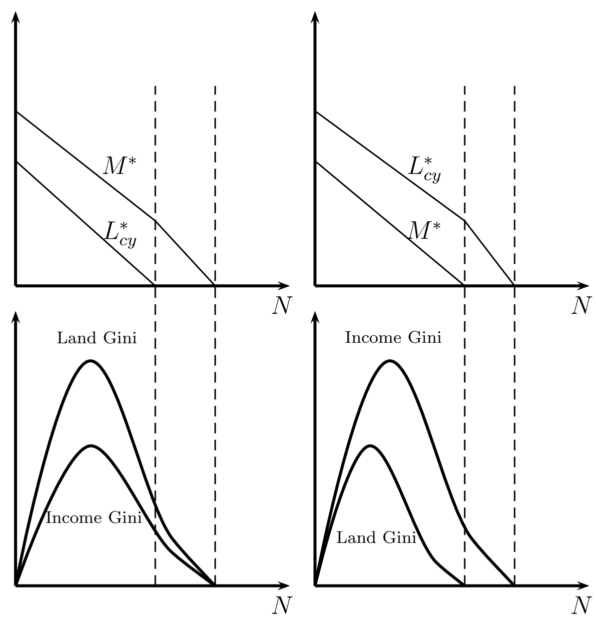

Figure 3: Factor intensity, optimal colonial investments, and the resulting land and income Gini coefficients of initial population density.

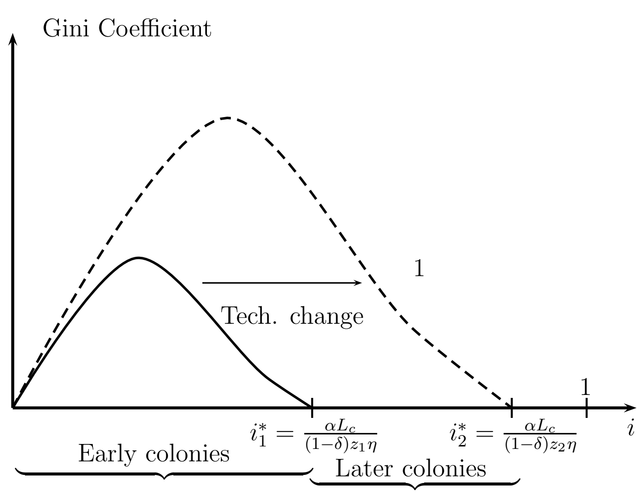

Figure 4: A Historical Interpretation of the Model with a Continuum of Colonies

Table 1: Descriptive statistics of main variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Gini index	83	47.421	9.168	28.3	64.9
Average population density per square km around 1500	83	5.275	12.276	0	100.46
Wheat-sugar ratio (in logs)	70	−0.011	0.143	−0.393	0.578
Latitude	83	0.178	0.12	0.011	9.667
Percent of land area in tropics	76	0.777	0.378	0	1
Distance from nearest coastline or navigable river in km	76	344.125	358.182	7.952	1466.67

Table 2: Gini index and population density circa 1500

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
sample	world	colonies	colonies	colonies no Egypt	colonies early tech	colonies new tech	colonies
lpopdens	-5.307***	-3.212***	3.409	7.746**	9.872***	-3.672***	-3.672***
	(0.79)	(0.92)	(3.44)	(3.75)	(3.70)	(1.03)	(1.03)
lpopdens2			-1.866**	-3.445***			
			(0.92)	(0.99)			
earlytech*lpopdens							13.54***
							(3.84)
earlytech							-7.369*
							(4.17)
Obs	130	83	83	82	29	54	83
R^2	0.29	0.11	0.16	0.19	0.25	0.16	0.25
F-stat	44.42***	11.77***	7.55***	34.94***	6.64**	12.17***	17.85***

Notes: All regressions are OLS with intercept (not shown). Dependent variable is inequality, measured by Gini index. Robust standard errors in parentheses. *, **, *** statistically significant at 10, 5, and 1 percent levels, respectively.

Table 3: "Horse race" of colonial determinants of inequality

	(1)	(2)	(3)	(4)	(5)	(6)
lpopdens1500	-4.061***	-3.593***	-3.966***	-4.108***	-4.155***	-5.090***
	(1.11)	(1.06)	(1.21)	(1.16)	(1.06)	(1.19)
earlytech*lpopdens	12.11**	11.10***	12.63***	14.07***	14.29***	13.82***
	(5.65)	(3.83)	(4.45)	(3.89)	(3.44)	(5.22)
earlytech	-6.121	-5.042	-6.983	-8.282*	-9.696**	-7.982
	(5.22)	(4.35)	(4.81)	(4.36)	(4.11)	(5.23)
lwheatsugar	-9.690					-5.121
	(8.04)					(10.1)
latitude		-18.31**				-43.84***
		(8.79)				(15.0)
tropics			3.272			-9.697*
			(3.21)			(5.64)
distcr				-0.0009		0.0013
				(0.004)		(0.003)
precious					3.882**	3.119
					(1.95)	(2.15)
Obs	70	83	76	76	83	69
R^2	0.31	0.30	0.28	0.27	0.28	0.42
F-stat	15.46***	18.14***	15.26***	13.96***	18.27***	24.27***

Notes: All regressions are OLS for ex-colonies sample (intercept not shown). Dependent variable is inequality measured by Gini index. Robust standard errors in parentheses. *, **, *** statistically significant at 10, 5, and 1 percent levels, respectively.

Table 4: Inequality in ex-colonies, excluding neo-Europes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
lpopdens	-4.775***	-5.379***	-4.713***	-5.554***	-5.462***	-5.138***	-6.386***
	(0.81)	(0.86)	(0.86)	(0.82)	(0.81)	(0.85)	(0.80)
earlytech*lpopdens	11.50***	12.29**	11.30***	13.07***	12.47***	12.72***	15.45***
	(3.97)	(5.36)	(4.05)	(4.30)	(4.05)	(3.90)	(5.42)
earlytech	-6.853	-7.116	-6.610	-8.705*	-8.390**	-9.262**	-11.54**
	(4.20)	(4.94)	(4.35)	(4.46)	(4.24)	(4.50)	(5.14)
lwheatsugar		-2.797					-5.858
		(7.76)					(9.72)
latitude			-3.878				-34.70**
			(8.25)				(16.2)
tropics				-1.363			-10.77**
				(2.22)			(4.99)
distcr					-0.000493		0.000205
					(0.0035)		(0.0034)
precious						2.986	3.203
						(1.97)	(2.15)
Obs	79	66	79	72	72	79	65
R^2	0.32	0.39	0.32	0.36	0.36	0.34	0.45
F-stat	24.96***	22.32***	18.20***	21.66***	21.38***	22.45***	18.87***

Notes: All regressions are OLS for ex-colonies sample without neo-Europes (Australia, Canada, NZ and USA). Dependent variable is inequality measured by the Gini index. Intercept not shown. Robust standard errors in parentheses. *, **, *** statistically significant at 10, 5, and 1 percent levels, respectively.

Table 5: Population density, inequality and income levels and growth

	(1)	(2)
Dependent variable	Instruments lpopdens & interaction	Instrument lpopdens
growth 1970-2000 (N=77)	**-0.029**	**-0.01**
r.s.e.	0.036	0.07
1st stage F-stat	14.72	7.98
Shea R^2	0.3	0.12
Anderson IV relevance test *p*-value	0.000	0.002
income 2000 (N=82)	**0.071*****	**0.142****
r.s.e.	0.025	0.068
1st stage F-stat	17.12	10.92
Shea R^2	0.24	0.10
Anderson IV relevance test *p*-value	0.000	0.003

Notes: All regressions are 2SLS (intercept not shown), with inequality measured by the Gini index. The growth regressions include the (log of) initial income as a basic control variable in the second stage.

Additional Figure and Tables for Top Quintile Income Share

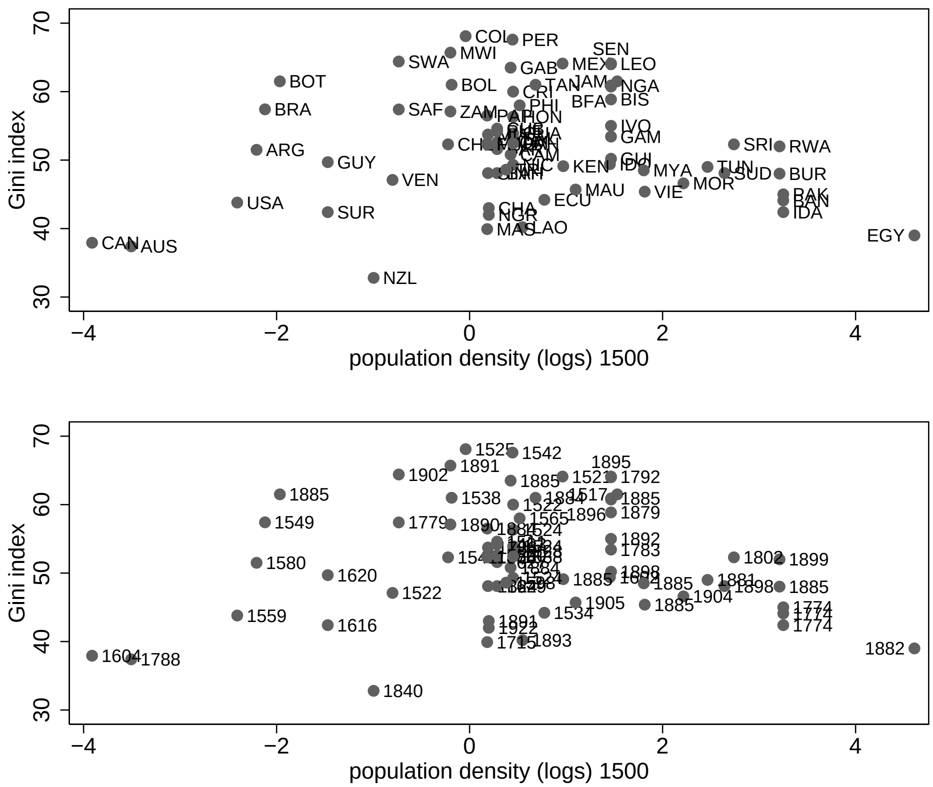

Figure 1: Scatter plot of (log) population density in 1500 and modern inequality

Notes: Inequality is measured by the top quintile income share in percent. Figure 2(a) above shows results listed by colony, Figure 2(b) below by year of colonization.

Table 1: Top quintile income share and population density circa 1500

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				colonies	colonies	colonies	
sample	world	colonies	colonies	no Egypt	early tech	new tech	colonies
lpopdens	-2.868***	-1.255	7.295**	9.908***	8.701***	-1.875	-1.875
	(0.70)	(0.96)	(3.04)	(3.74)	(2.86)	(1.16)	(1.16)
lpopdens2			-2.3***	-3.217***			
			(0.72)	(0.97)			
earlytech*lpopdens							10.58***
							(3.09)
earlytech							-7.175*
							(3.71)
Obs	103	69	69	68	27	42	69
R^2	0.14	0.02	0.15	0.13	0.23	0.06	0.13
F-stat	16.33***	1.65	8.92***	10.28***	8.55***	2.48	5.36***

Notes: All regressions are OLS (intercept not shown). Dependent variable is inequality, measured by income share of top quintile. Robust standard errors in parentheses. *, **, *** statistically significant at 10, 5, and 1 percent levels, respectively.

Table 2: "Horse race" of colonial determinants of inequality measured by top quintile income share

	(1)	(2)	(3)	(4)	(5)	(6)
lpopdens	-2.105*	-1.806*	-1.935	-2.214*	-2.497**	-2.743**
	(1.19)	(1.00)	(1.20)	(1.26)	(1.24)	(1.24)
earlytech*lpopdens	8.383*	8.027**	8.588**	10.31***	12.03***	9.595*
	(4.66)	(3.34)	(3.71)	(3.11)	(2.97)	(5.41)
earlytech	-6.591	-4.886	-6.226	-7.811**	-10.37***	-7.809
	(4.41)	(3.80)	(4.19)	(3.76)	(3.81)	(4.86)
lwheatsugar	-12.44					-3.403
	(8.48)					(11.5)
latitude		-21.69***				-30.75*
		(7.58)				(16.3)
tropics			5.499*			-4.451
			(3.13)			(6.25)
distcr				-0.00220		-0.000391
				(0.0029)		(0.0030)
precious					4.097**	2.110
					(1.86)	(2.00)
Obs	60	69	64	64	69	59
R^2	0.18	0.23	0.21	0.16	0.18	0.29
F-stat	7.22***	7.9***	7.80***	4.39***	4.82***	6.15***

Notes: All regressions are OLS for ex-colonies sample (intercept not shown). Dependent variable is inequality, measured by income share of top quintile. Robust standard errors in parentheses. *, **, *** statistically significant at 10, 5, and 1 percent levels, respectively.

Table 3: Top quintile income share in ex-colonies - excluding neo-Europes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
lpopdens	-3.184***	-3.677***	-3.029***	-3.709***	-3.773***	-3.681***	-4.399***
	(0.81)	(0.77)	(0.83)	(0.79)	(0.77)	(0.96)	(0.87)
earlytech*lpopdens	8.942***	9.757**	8.875***	9.584***	9.351***	11.23***	12.60**
	(3.05)	(4.07)	(3.29)	(3.20)	(3.16)	(3.65)	(5.85)
earlytech	-7.402**	-8.501**	-7.233**	-8.644**	-9.093***	-11.05**	-13.10***
	(3.37)	(3.63)	(3.55)	(3.45)	(3.34)	(4.37)	(4.90)
lwheatsugar		-0.822					-1.854
		(6.96)					(11.2)
latitude			-8.863				-25.27
			(7.70)				(20.6)
tropics				0.327			-6.672
				(1.84)			(5.24)
distcr					-0.00224		-0.00222
					(0.0033)		(0.0033)
precious						3.307	2.592
						(2.10)	(2.23)
Obs	65	56	65	60	60	65	55
R^2	0.18	0.24	0.19	0.23	0.24	0.22	0.30
F-stat	8.85***	9.44***	7.47***	8.91***	9.05***	5.87***	4.86***

Notes: All regressions are OLS for ex-colonies sample without neo-Europes (Australia, Canada, NZ and USA). Dependent variable is inequality, measured by top quintile income share. Intercept not shown. Robust standard errors in parentheses. *, **, *** statistically significant at 10, 5, and 1 percent levels, respectively.

Table 4: Population density, inequality and income levels and growth

	(1)	(2)
Dependent variable	Instruments lpopdens & interaction	Instrument lpopdens
growth 1970-2000 (N=65)	-0.048	-0.025
r.s.e.	0.047	0.117
1st stage F-stat	5.77	3.69
Shea R^2	0.21	0.06
IV relevance test P-value	0.019	0.103
income 2000 (N=68)	0.065	0.396
r.s.e.	0.05	0.4
1st stage F-stat	5.23	1.54
Shea R^2	0.13	0.02
IV relevance test P-value	0.052	0.288

Notes: All regressions are 2SLS (intercept not shown), with inequality measured by the the top quintile income share. The growth regressions include the (log of) initial income as a basic control variable in the second stage.

Working Papers of the Center of Economic Research at ETH Zurich

(PDF-files of the Working Papers can be downloaded at www.cer.ethz.ch/research).

08/86 M. J. Baker, C. N. Brunnschweiler, and E. H. Bulte
Did History Breed Inequality? Colonial Factor Endowments and Modern Income Distribution

08/85 U. von Arx and A. Ziegler
The Effect of CSR on Stock Performance: New Evidence for the USA and Europe

08/84 H. Gersbach and V. Hahn
Forward Guidance for Monetary Policy: Is It Desirable?

08/83 I. A. MacKenzie
On the Sequential Choice of Tradable Permit Allocations

08/82 I. A. MacKenzie, N. Hanley and T. Kornienko
A Permit Allocation Contest for a Tradable Pollution Permit Market

08/81 D. Schiess and R. Wehrli
The Calm Before the Storm? - Anticipating the Arrival of General Purpose Technologies

08/80 D. S. Damianov and J. G. Becker
Auctions with Variable Supply: Uniform Price versus Discriminatory

08/79 H. Gersbach, M. T. Schneider and O. Schneller
On the Design of Basic-Research Policy

08/78 C. N. Brunnschweiler and E. H. Bulte
Natural Resources and Violent Conflict: Resource Abundance, Dependence and the Onset of Civil Wars

07/77 A. Schäfer, S. Valente
Habit Formation, Dynastic Altruism, and Population Dynamics

07/76 R. Winkler
Why do ICDPs fail? The relationship between subsistence farming, poaching and ecotourism in wildlife and habitat conservation

07/75 S. Valente
International Status Seeking, Trade, and Growth Leadership

07/74 J. Durieu, H. Haller, N. Querou and P. Solal
Ordinal Games

07/73 V. Hahn
Information Acquisition by Price-Setters and Monetary Policy

07/72 H. Gersbach and H. Haller
Hierarchical Trade and Endogenous Price Distortions

07/71 C. Heinzel and R. Winkler
The Role of Environmental and Technology Policies in the Transition to a Low-carbon Energy Industry

07/70 T. Fahrenberger and H. Gersbach
Minority Voting and Long-term Decisions

07/69 H. Gersbach and R. Winkler
On the Design of Global Refunding and Climate Change

07/68 S. Valente
Human Capital, Resource Constraints and Intergenerational Fairness

07/67 O. Grimm and S. Ried
Macroeconomic Policy in a Heterogeneous Monetary Union

07/66 O. Grimm
Fiscal Discipline and Stability under Currency Board Systems

07/65 M. T. Schneider
Knowledge Codification and Endogenous Growth

07/64 T. Fahrenberger and H. Gersbach
Legislative Process with Open Rules

07/63 U. von Arx and A. Schäfer
The Influence of Pension Funds on Corporate Governance

07/62 H. Gersbach
The Global Refunding System and Climate Change

06/61 C. N. Brunnschweiler and E. H. Bulte
The Resource Curse Revisited and Revised: A Tale of Paradoxes and Red Herrings

06/60 R. Winkler
Now or Never: Environmental Protection under Hyperbolic Discounting

06/59 U. Brandt-Pollmann, R. Winkler, S. Sager, U. Moslener and J.P. Schlöder
Numerical Solution of Optimal Control Problems with Constant Control Delays

06/58 F. Mühe
Vote Buying and the Education of a Society

06/57 C. Bell and H. Gersbach
Growth and Enduring Epidemic Diseases

06/56 H. Gersbach and M. Müller
Elections, Contracts and Markets

06/55 S. Valente
Intergenerational Transfers, Lifetime Welfare and Resource Preservation

06/54 H. Fehr-Duda, M. Schürer and R. Schubert
What Determines the Shape of the Probability Weighting Function?

06/53 S. Valente
Trade, Envy and Growth: International Status Seeking in a Two-Country World

06/52 K. Pittel
A Kuznets Curve for Recycling

06/51 C. N. Brunnschweiler
Cursing the blessings? Natural resource abundance, institutions, and economic growth

06/50 C. Di Maria and S. Valente
The Direction of Technical Change in Capital-Resource Economics

06/49 C. N. Brunnschweiler
Financing the alternative: renewable energy in developing and transition countries

06/48 S. Valente
Notes on Habit Formation and Socially Optimal Growth

06/47 L. Bretschger
Energy Prices, Growth, and the Channels in Between: Theory and Evidence

06/46 M. Schularick and T.M. Steger
Does Financial Integration Spur Economic Growth? New Evidence from the First Era of Financial Globalization

05/45 U. von Arx
Principle guided investing: The use of negative screens and its implications for green investors

05/44 Ch. Bjørnskov, A. Dreher and J.A.V. Fischer
The bigger the better? Evidence of the effect of government size on life satisfaction around the world

05/43 L. Bretschger
Taxes, Mobile Capital, and Economic Dynamics in a Globalising World

05/42 S. Smulders, L. Bretschger and H. Egli
Economic Growth and the Diffusion of Clean Technologies: Explaining Environmental Kuznets Curves

05/41 S. Valente
Tax Policy and Human Capital Formation with Public Investment in Education

05/40 T.M. Steger and L. Bretschger
Globalization, the Volatility of Intermediate Goods Prices and Economic Growth

05/39 H. Egli
A New Approach to Pollution Modelling in Models of the Environmental Kuznets Curve