European Bank
for Reconstruction and Development

Labour markets and representative institutions: evidence from colonial British America

Elena Nikolova

Summary

The literature has identified the quality of political institutions in the transition region as an essential but under-studied component of growth, transition and reform. This paper aims to disentangle the determinants of democratic institutions by investigating if high income inequality is always detrimental to the emergence and stability of such regimes, and under what circumstances labour scarcity can mitigate its adverse effect. Instead of using a cross-country data set of transition countries covering only 20 years and prone to simultaneity and reverse causality, I present a quantitative analysis of suffrage restrictions in the 13 British American colonies from their establishment to the American Revolution. I show that in cases of a labour shortage elites use the right to vote as a tool to attract workers. Democratic institutions thus emerge despite high inequality and redistribution pressures, and only when labour demand eases do those in power have an incentive to contract the suffrage, as predicted by standard inequality-democracy arguments. Using a new and unique colony-level panel data set covering nearly 150 years, I estimate that a 10 percentage point decrease in the scarcity of labour, conditional on inequality and other controls, decreases the inclusiveness of political institutions by over 8 per cent in the short run, and by close to 20 per cent in the long run.

Working Paper No. 134   **Prepared in October 2011**

1 Introduction

Why have transition countries such as Poland and the Czech Republic been able to grow faster than others, like Tajikistan and Kazakhstan? Researchers have acknowledged that purely economic explanations, such as those focusing on macroeconomic polices and reform strategies, are insufficient. Instead, institutions in general, and political institutions in particular, have been shown to affect growth, privatisation and reform in the transition region.[1] However, few studies have explored why transition countries differ so widely in the quality of their political institutions in the first place. This is not only because the 20 years that have elapsed since the fall of communism have not witnessed significant regime changes, but also because utilising a cross-country data set of transition countries is subject to data inconsistencies and unit heterogeneity as well as to econometric issues relating to variable endogeneity and reverse causality. This paper seeks to draw causal inferences for the evolution of democracy in the transition region by exploiting the case study of the 13 British American colonies from their establishment to the American Revolution.

While at first glance the economic history of an advanced democracy may seem irrelevant for the transition region, there are at least two parallels between the American colonies and states and the countries that were part of the former Soviet bloc. First, American colonial and post-independence development was affected by forces similar to those that have brought increased economic and political interdependence today. For instance, exogenous shocks such as slavery and industrialisation redefined labour relations and institutional arrangements in the South, just as in some countries political power and redistribution have been affected by global trade fluctuations and the economic crisis. Similarly, the Southern colonies and states and some of today's transition countries, such as those in Central Asia, share a dependence on agriculture and commodity exports.

Recent research on the causes of democracy has linked the probability of democratic transition and consolidation to low levels of inequality, relatively mobile wealth, economic development and historical path dependence.[2] More specifically, part of this new research has built on variations of the classical Meltzer-Richard model of taxation and redistribution to claim that places with a highly unequal income distribution are less likely to be democratic, as the rich prefer an authoritarian regime due to redistribution pressures from a low-income median voter. This paper argues that although such models are important, they are rather incomplete, as the structure of labour markets often trumps inequality in the final choice of the franchise regime.

In particular, I show that liberal representative institutions will arise when there is a labour shortage and high inequality. On the one hand, an inclusive suffrage increases elite profits by attracting poor labourers, relative to the loss of redistribution arising from high inequality. On the

---

[1]For studies related to the transition region, see Roland (2002) and Beck and Laeven (2006). For a general treatment of the relationship between democracy and economic development, see Persson and Tabellini (2006).

[2]Acemoglu and Robinson (2005), Lizzeri and Persico (2004) and Engerman and Sokoloff (2000, 2002 and 2005) explore the link between inequality, redistribution and regime choice; Boix (2003) studies how inequality and the asset specificity of wealth affect democratic emergence and transition. On democracy and development, see Lipset (1959), Przeworski and Limongi (1997) and, more recently, Boix (2010) who finds that the democracy-development relationship is robust in the cross-country historical context. On how critical junctures affect democratic institutions, see Acemoglu et al (2008b).

other hand, when there is a labour surplus, those with the power to limit the suffrage will do so even when income is relatively equally distributed, as democracy would otherwise increase the number of workers in an already satiated labour market. In the novel model of this paper, the extensiveness of the suffrage is not only the usual determinant of equilibrium taxation and redistribution, but also a tool for profit maximising elites to credibly commit to wage and redistribution promises which they made to poor workers over the long run.[3] While having the right to vote allows the poor to get the highest possible short-term redistribution, a representative assembly dominated by those with low incomes also serves to discipline any rich members wishing to implement measures aimed at reducing the long-term welfare of the poor. Therefore, a place with high wages *and* a wide suffrage is particularly attractive for labourers with an income below the mean.

Conversely, the extensiveness of the suffrage can affect the welfare of the rich in two opposite ways. On one hand, strict voting restrictions enable the rich to control the assembly, which allows them to increase their incomes by passing laws aimed at depressing redistribution and the compensation paid to poor workers. On the other hand, a liberal suffrage attracts poor workers, which increases the amount of resources produced in the economy and elite profits as long as the production function is upward sloping (as will be the case when scarce labour is the primary input in production). How elites weigh these two payoffs depends on whether labour markets are tight or easy and determines what type of institutions will ultimately emerge.

I apply these theoretical insights to the political development of the 13 British American colonies from the early 17th century until the American Revolution. In order to do so, I use a new and unique colony-level data set which provides annual observations for a wide range of political, economic and social variables, including suffrage, labour markets, inequality and economic development. I constructed the data set using a large number of secondary data sources, which were cross-checked whenever possible and which, to the best of my knowledge, have not been used before in the literature. The data show that in the South, the tropical climate and less favorable disease environment discouraged British immigration in the early years of settlement. Since the colonial economy was largely agricultural, the wealthy landowners, who depended on the labour of the poor, established liberal representative institutions to attract migrants. An inclusive political system guaranteed workers that the fruits of their labour would not be expropriated and prevented the rich from reneging on their promises of land and wages. At the same time, Southern elites did not always ignore the income distribution when altering representative institutions: as rising inequality in the last quarter of the 17th century made extending the suffrage less profitable, voting restrictions were reintroduced (although they still remained lower than those in the North). Suffrage conditions became restrictive once rich planters had access to cheap slave labour and thus did not need to grant any political concessions to poor white voters.

Instead, the North appealed to migrant families who wanted to settle permanently in the mainland colonies. Since the climate of the region was more temperate and its soils could only support small family-based farms, population increase through reproduction was sufficient to satisfy labour demand. As a result, relatively stricter suffrage regulations in the pre-1700 North (such as a religious restriction) were introduced to curb any additional immigration that could not be

---

[3]In the 13 colonies, assembly representatives could easily control wages or the dues payable to indentured labourers at the end of their term by passing the appropriate laws, such as minimum wage laws (Purvis, 1999).

supported given the limited opportunities to grow export crops. The slavery shock had a limited impact on labour markets in the North, and Northern production continued on small family farms with an adequate supply of labour. As a result, political institutions remained largely unchanged.

In order to measure the extensiveness of the right to vote, I create an index that aggregates various restrictions on the franchise in each colony, such as those related to income, land ownership or religion. Labour markets are proxied by a colony's percentage of the black population: a relatively large black population indicates a labour market surplus (as in the South after 1700), whereas a low value of this variable indicates a labour shortage (as in the South before 1700) or just enough labour (as in the North). I use the percentage of a colony's landless population to estimate inequality, and argue that a higher percentage landless implies more inequality, and vice versa. Finally, I capture levels of economic development through a wage index, as economic theory predicts that real wages should be higher in richer places.

I find that labour market tightness has a strong impact on suffrage (conditional on inequality, economic development and a rich set of controls). A 10 percentage point increase in the supply of labour increases the strictness of the suffrage by over 8 per cent in the short run, and by close to 20 per cent in the long run, relative to the mean. This effect is robust to various regression specifications, the inclusion of numerous additional controls, and persists after the variable capturing labour markets is instrumented with a crop index measuring each colony's propensity to produce tobacco, wheat or rice, interacted with the price of the respective crop. Neither do the results change when I experiment with two additional measures of colonial labour markets: Caribbean slave prices, scaled by the crop index variable; and a binary variable that is 1 for the South after the arrival of slavery in 1700, and 0 otherwise.

This paper aims to remedy at least two flaws found in the existing work on representative institutions. The most influential studies specify broad causal mechanisms and then proceed to test the theories on a sample of countries as diverse as Switzerland and Somalia. However, such an approach is theoretically weak, as it ignores factors that are specific either to a particular period or country. Second, from an empirical point of view, using cross-country data is prone to identification issues arising from data inconsistency, unit heterogeneity and endogeneity.[4] In addition, these papers typically lack country fixed effects and employ dubious instruments to resolve endogeneity and reverse causality.[5]

The paper addresses these methodological shortcomings by exploiting the wide variation in the suffrage restrictions of the 13 British American colonies from their initial establishment to the American Revolution. My novel research setup examines a group of similar colonies settled around the same time and exploits the exogenous nature of geographic characteristics and the slavery shock.[6] Although the research design is not an historical experiment because the colonies

---

[4]For a discussion of the theoretical and empirical problems of cross-country data, which are used by most of the works cited in footnote 1, see Pande and Udry (2005) and Atkinson and Brandolini (2001). Country-level institutional studies in the spirit of this paper include Jha (2008), Banerjee and Iyer (2005), Nunn (2008) and Acemoglu et al (2008a).

[5]For example, Albouy (2008) presents a critique of the use of settler mortality as an instrument for institutions.

[6]In essence, the research setup incorporates two quasi-experiments: one is the initial settlement of the colonies in the early 17th century, while the second one is the slavery shock in the early 18th century. While focusing on a group of similar colonies eliminates the problem of cross-country heterogeneity, colony heterogeneity may be a

differed from the very beginning on the dependent variable and because suffrage restrictions were continuously affected by the state of labour markets, the design does allow to disentangle the effect of labour markets from that of other variables correlated with inequality or economic development, such as geography or ethno-linguistic fractionalisation.[7]

This paper is organised as follows. Section 2 provides a brief historical background, and Section 3 develops the theoretical argument. Sections 4 and 5 outline the nature of the data and the empirical setup. Section 6 discusses the quantitative results, while Section 7 focuses on alternative mechanisms. Section 8 concludes. The Supplementary Materials present the formal model and the results from regression robustness checks.

2 Historical background

2.1 General overview

The primary goal of American settlement was to remedy the twin problems of scarcity of land and abundance of labour facing Britain in the 17th century. The British mainland colonies followed three distinct patterns of settlement: charter, proprietary and royal. Charter colonies originated when the Crown granted patents and a land grant to a corporation or trustees without imposing any royal supervision. In proprietary colonies, one or more Britishmen gained ownership of a colony, on the condition that the granted area became settled with British citizens. The government encouraged settling the colony at private expense, and the official endorsement aided the influx of migrants who bought land from the proprietor and then paid him quitrents. Since the proprietary arrangement worked to the advantage of both the government and the colony owner, most of the 13 colonies were settled through this system. Finally, royal colonies were governed directly by Britain through an appointed governor.

At least two-thirds of the colonists were engaged in agriculture, and the top three export crops were tobacco, rice and wheat (Purvis 1999, 50). The geographical distribution of the three plants reflected their biological characteristics: tobacco was grown mainly in the Upper South, rice in the swampy areas of the Lower South, and wheat in New England (in the 17th century) and in the Middle colonies (starting in the late 1600s) (Purvis 1999, 76).[8] These crops exhibited different economies of scale in production. Rice was most efficiently grown with slave labour not only because whites could not tolerate the disease-ridden swampy areas of the rice fields, but also because rice production could be divided into a series of "assembly-line" tasks (Friedman 1989, 194). However, not all operations involved in growing tobacco could be done gang-style: some required special skills and needed to be performed by only a few labourers.[9] Finally, large slave

---

problem, but to a lesser extent as compared with the cross-country setting. Time invariant and time-variant colony-specific differences can be easily handled by including geographic and time controls, as I do later in the paper.

[7]On the importance of geography, see Bloom and Sachs (1998), and on the link between ethnic fractionalisation and growth/development see Easterly and Levine (1997) and Miguel (2004).

[8]Wheat is a less capricious plant compared with rice, indigo or tobacco, and can be grown virtually anywhere (email communication with Todd Ballard, April 13, 2009). Therefore, as tobacco growing became less profitable in the late 18th century, Upper South farmers allocated more of their resources to wheat cultivation (see Klingaman 1969).

[9]Fogel (1989, note 53, 428) gives an example of two different types of tobacco tasks. The tobacco replanting

plantations were completely inappropriate for growing wheat: the labour needs for grain were seasonal and could not employ slaves or indentured servants during the whole year (Purvis 1999, 78).

Although all 13 colonies were under British rule, they were essentially self-governing units. Even in the royal colonies the elective assembly controlled colonial finances (including the governor's salary) and the power of the Crown was significantly curtailed (Purvis 1999, Galenson 1996). As a result, there was substantive variation in colonial suffrage restrictions (Purvis 1999, 187). While some colonies reproduced exactly the British 40 shilling freehold requirement, other British rules were not directly adopted and colonists were free to choose the suffrage regulations that best suited them without any interference from the Crown. Furthermore, colonial records indicate that suffrage restrictions were binding, suggesting that the voting laws that were on paper should be highly correlated with those that were actually enforced (Keyssar 2000; Dinkin 1977, 28-29, 47).

Three broad colonial social classes emerged based on the distribution of income: the rich, the middle class and the poor. In the North and Middle, the rich were merchants trading with Britain (Brown 1969, Nash 1979), whereas in the South elites included mostly large planters.[10] Small and medium freeholders (or the middle class) made up a large proportion of the population throughout the colonial period in the North and Middle colonies.[11] Although indentured servants were initially the largest group in the South, small and medium farmers became more numerous after the introduction of slavery.[12] Finally, the poor were indentured servants and whites without land or slaves (the latter population usually consisted of newly freed servants). Indentured servants were British immigrants that committed themselves to working for a landholder in order to obtain money for travelling across the Atlantic. While in the North indentured servants were a very small percentage of the population, in the South they were numerous. For example, in 17th century Maryland, this group constituted 70 to 85 per cent of all immigrants (Jordan 1987, 13). The poor dominated the population in the South until the arrival of slavery and the tightening of the suffrage at the beginning of the 18th century prompted them to migrate to other colonies or to attempt to obtain slaves at all costs.

2.2 Evolution of representative institutions in the 13 colonies

Chart 1 summarises graphically how suffrage restrictions (*Suffrage*) changed in the South and North.[13] When designing colonial representative institutions, elites weighed the gains from at-

---

procedure was done about six weeks after the seeds were planted. The withdrawal of the plants from their beds required handicraft skills and was therefore not performed by large slave groups, but gangs were involved in the actual replanting of the plants.

[10]Perkins (1980, 58) points out that full planter status came with the possession of 500 acres of land and 20 slaves.

[11]In most colonies, a freeman was a free male inhabitant (in the early years of the charter colonies, a freeman signified a member of the colonial corporation). A freeholder was a freeman who owned land.

[12]When characterising this class in the Southern context, I follow the definition provided by Williams (1969) for Virginia: a small farmer owned between 50 and 500 acres of land. An easier approach would be to count any farmer owning less than 10-20 slaves as a member of the middle class, but the fact some planters in the South owned slaves but rented the land makes such a definition problematic.

[13]A higher value of Suffrage is equivalent to less democratic institutions, and vice versa. See Section 4 and Appendix 1 for information about how this variable was constructed.

tracting migrants from Britain versus the losses from having to stick to previous promises of wages and land grants. While before 1700 Southern elites opted for liberal representative institutions in order to encourage British immigration, stricter voting regulations in the North were used to limit the number of settlers (see Morgan 1975; Purvis 1999).[14] At the same time, suffrage restrictions in the 17th century South were on the rise but remained below those in the North. Even though the South experienced a shortage of labour throughout the 17th century, rising life expectancy and higher inequality led to a gradual contraction of the right to vote.

The arrival of slavery prompted Southern elites to tighten the suffrage even further, as a steady supply of slaves meant that attracting British migrants was no longer necessary. Instead of devising policies catering to poor workers from Britain, colonial assemblies were now focused on issues concerning only landholders, such as tobacco or rice price fluctuations, land expropriation threats from Britain, as well as debt relations between Southern planters and British gentlemen. Since the rich did not have to spend their increasing wealth on compensating the "losers" from the labour market shock (landless whites) and colonial funds could be allocated to issues of interest only to landholders, suffrage was limited to the class of freeholders.

In the North, unlike in the South, the main commodities (grains, potash, fish, whale products and livestock) did not exhibit economies of scale, and no additional supply of labour was needed both before and after the shock of slavery (McCusker and Menard 1991). In the New England colonies, the balance between the sexes allowed the population to grow without the need for additional migration, while the Middle colonies also welcomed immigrants from Europe, who settled mainly in the cities after 1700. Northern elites kept suffrage restrictions at the level where the number of workers satisfied labour demand, and since labour markets were not affected by slavery or any other factors, representative institutions remained unchanged too.[15]

3 Theory

3.1 Specification of the causal mechanisms

Following the workhorse model of political decision-making (Meltzer and Richard 1981), assume a simple economy populated by rich and poor agents who do not exhibit any within-class differences and where, consistent with the historical evidence discussed in the following section, the median voter is always a poor individual (for a formal setup and proofs, see the Supplementary Materials).[16] Then, as inequality increases, the rich are more likely to implement an autocratic

---

[14]Galenson (1984) points out that the liberal right to vote was underpinned by material work incentives such as giving indentured servants food, lodging and clothing above the minimum amount required by law, or even voluntarily increasing the freedom dues given to bound labourers.

[15]Note that Chart 1 shows a slight decrease in Northern suffrage restrictions after 1700, with the inflection point around 1750. Given that the French and Indian war took place from 1754-1763 and involved New England and the Middle colonies, a potential explanation could be that Northern suffrage decreased in order to recruit soldiers. On this topic, see Ticchi and Vindigni (2008) and Przeworski (2007).

[16]While the historical section shows that the 13 colonies also had a middle class, the current simplification does not change the nature of the results as the median voter was a poor individual throughout the colonial period. The existence of a middle class may render it more profitable for the rich to impose a partial democracy instead of an autocracy in the case of a labour surplus if such an alignment allows the rich to counteract any security threats or to lobby more effectively the British governmnent (see the Supplementary Materials).

system, and vice versa. I adopt a broad definition of "redistribution": it encompasses all returns of a particular group when given the right to vote. As an example, controlling the Assembly allows the poor to increase their income not only through direct redistribution, but also through the control of wages and the use of public resources and policies, such as special financing.

Compared with the standard model, this political economy has the following additional traits. First, output is produced according to a Cobb-Douglas production function with only land and labour (the supply of land is exogenous and fixed). Only the poor, who are initially in another country, are labourers, while the rich landholders make all economic and political decisions, including those relating to wages, any redistribution or other concessions to the poor, and the extent of the suffrage. In order to attract labour to the colonies, the rich offer foreign workers the highest possible wage - that of a competitive labour market. However, these promises are not credible, as elites would prefer to pay lower wages once workers arrive. This setup assumes that at least some of the workers have high re-entry costs once they arrive in the colonies. The quick substitution of poor workers with slaves after 1700 shows that rich landowners and colonial companies would have preferred to capture the highest possible share of output even when labour was scarce.

To convince labourers that any promises made in order to attract them to the colonies will be kept, elites also need to extend the right to vote to the poor. Furthermore, regime choice has two opposing choices on the welfare of the rich: democracy raises taxes relative to autocracy but also attracts workers to the colonies.[17] For the poor, democratic institutions bring both competitive-market wages and redistribution because of higher taxation, while an autocracy only gives them a subsistence wage without any other concessions or redistribution. Moreover, the income equation of the poor also includes two additional terms: positive non-monetary benefits of voting (when democratic institutions prevail) and a negative cost of migrating to the foreign country.

Consider how the model works over a two-period game. In period 0, only the rich have the right to vote. In the first period, the rich grant the suffrage to the poor or keep it restricted, and after that the poor choose whether to migrate based on a comparison of their net payoffs in the two countries. In the second period, an exogenous labour shock occurs: equally productive slaves who do not demand any political rights and are only paid a subsistence wage are available to the rich, with some probability. The rich make the decision about political institutions based on their expectation of labour supply, conditional on inequality.[18] After the political system is announced, the poor decide whether to stay or to return to the home country (or go to another colony).

Consequently, when choosing whether to leave their home country in the first period, the poor know that because of credibility issues, the rich can only be held accountable to their promises of wages and taxation if inclusive representative institutions exist. There is a range of parameter assumptions under which a large number of labourers choose to come to the new country in the first period only if the residing elites adopt a democratic political system. In the second period, all poor will stay if liberal representative institutions are kept, while some of them will leave for

---

[17]The functional specification not only allows output (and the income of the rich) to directly depend on the labour supply, but also provides a simple measure of inequality as the share of income accruing to the poor.

[18]Total labour supply is also affected by the proportion of poor workers that decide to stay after a contraction of the suffrage due to high moving costs.

a place with better voting rights when autocracy is implemented.[19] For elites, the relative size of the poor population generates two opposing forces which ultimately determine what type of political system will emerge. When deciding whether to extend voting rights to the poor, elites face a trade-off between more redistribution and attracting more workers. While redistribution always decreases the income of the rich, a higher supply of labour will either increase elite profits when there is a labour shortage, or not affect them when there is a labour surplus.

In other words, in contrast to traditional democracy-inequality arguments predicting that high inequality always leads to a contraction of the suffrage because of redistribution pressures, elites will have an incentive to extend the suffrage if attracting labour raises profits in cases of a tight income distribution *and* a labour shortage. However, once labour pressures subside, the rich no longer need to keep liberal institutions, and political rights are taken away, even in cases of low inequality.

3.2 Suffrage institutions and migration

Were liberal suffrage institutions effective in attracting poor British migrants to the South? To answer this question, consider first the existing (fragmentary) data on the estimated British immigration to the New World covering the colonies of Virginia, Maryland, New England, and Delaware (Purvis 1999, 164). Consistent with the North-South franchise difference before the arrival of slavery, during the period 1607-1700, the number of migrants (per square mile of settled territory) coming to Virginia and Maryland was approximately double that choosing New England and Delaware.[20] While there are no direct data comparing the migration flows of the South with those of the North after 1700, data from the colonial militia muster rolls (covering the French and Indian war and the pre-revolutionary years) analysed by Sokoloff and Villaflor (1982) show an immigration trend different from the one observed in the 17th century. While the South drew a lot of migrants before slavery, it was no longer the "top destination" for migrants after 1700: the Carolinas, New England and New York were equally attractive to native and foreign workers. What is more, Maryland and Virginia were even losing American-born men (Virginia experienced a net loss even after counting foreign immigration).

The colonial distribution of the foreign-born, which would be a good proxy for migration from Europe and Britain, suggests a similar picture after 1700. While the foreign-born population of the Middle colonies was rather high (74 per cent of total population in Pennsylvania and 61 per cent of total population in Delaware), the proportions were very similar in the South and New England: 20 per cent in Virginia, 16 per cent in South Carolina, 9 per cent in North Carolina, versus 17 per cent in Massachusetts and 10 per cent in New Hampshire. The only outlier is

---

[19]Those poor for whom reverse migration is too costly could react to a tightening of the right to vote by rebelling, but this action is not included in the model. In the current model, the poor who have high re-entry costs simply stay in the colonies after a tightening of the suffrage.

[20]For settled territory calculations, see Purvis (1999), 19 and Purvis (1994), 12 and 243-244. While far from conclusive, data presented in Dunn (2000), 55 shows similar trends. In 1635, 2,013 London immigrants chose to go to Virginia, 1,177 to New England, and 1,700 to the Island colonies. Note also that the West Indies attracted the largest number of British immigrants before 1607-1700 - a total of 212,400. The reason for this is that tobacco and sugar production started much earlier on the islands relative to the mainland colonies. However, West Indian planters adopted slavery in the 1660s and Virginia became the most desirable destination for British indentured servants (Dunn 2000, 70).

Maryland, with the foreign born constituting 56 per cent of its population.[21]

Some indirect evidence suggests that Southern representative institutions were indeed used as a tool to attract workers. Scholars agree that colonies with labour shortages initially had a high proportion of men, while those with enough labour supply had a relatively balanced sex ratio (Purvis 1999). Due to the hotter climates and more virulent disease environment, the Southern colonies at first had a very small female population, and therefore producing a steady labour supply or attracting whole families was difficult.[22] However, because the climate of the Northern colonies was similar to that in Britain, immigrants brought with them their families, and the natural growth rate of the population satisfied labour demand. Therefore, if lower suffrage restrictions were used to attract migrants, there should be a negative correlation between Suffrage and the Sex Ratio of men to women (*Sex Ratio*) across the 13 colonies. Chart 2 confirms this relationship. Similarly, Suffrage and *Life Expectancy* in the South are positively correlated (Chart S2 in the Supplementary Materials).

The idea that the elites' need for settlers determines the evolution of institutions has parallels in other work. Bertocchi and Strozzi (2008) suggest that the quality of electoral institutions was important in attracting European migrants to the New World colonies in the 19th century. Although the time period at which the authors look is later than that under consideration in this paper, the same logic can be applied to colonial migration. Likewise, Wright (2006) explains that slavery was important for Southern economic development not because slaves were more productive than white labourers, but because planters could force slaves to work in places where it would have been difficult to attract free labour.[23]

4 Data summary

I measure the extensiveness of the suffrage using several related variables: the sum of indicators capturing all suffrage restrictions listed in McKinley (1905) (Suffrage), the sum of only the suffrage restrictions identified by Dinkin (1977) as important (*Suffrage Dinkin*), and a restriction index obtained by principal component analysis on Suffrage.[24]

Because institutions change slowly, the theory presented in this paper is likely to explain only medium- and long-term changes in the dependent variable. Therefore, instead of using an annual

---

[21]The article points out that foreign migrants chose to arrive in urban centers, and the rise of Baltimore as well as the proximity of Maryland to the Middle colonies could account for the high fraction of foreign born population in this colony.

[22]Purvis (1999, 153) argues that by the 1700s, the slow increase of Southern households eventually decreased the relative imbalance of men versus women. See Edlund (2005) who documents that rural areas in the Western world are relatively short on women, compared with urban areas. Basher and Lagerlöf (2008) show that there is a negative correlation between the sex ratio and log population density among the US states in 1900.

[23]The argument of the paper is also similar to industrial organisation studies arguing that buyers may be willing to grant extra concessions to sellers on whom they depend (see Grossman and Hart 1986 and Hart and Moore 2007).

[24]For more on the coding procedures and what types of restrictions are included in the dependent variable, see Appendix 1. Using principal component analysis aims to give data-dependent weights for each suffrage restriction instead of weighing voting regulations equally. An alternative way of determining the weights would be to consider how much each restriction contributed to turnout in the colonial period, but such data are too fragmentary to allow a consistent analysis.

panel for the regression analyses, I create unbalanced 5-year and 10-year panels by taking every fifth and tenth observation for each colony.[25] I drop observations for the Middle colonies during Dutch or Swedish rule in order to avoid any confounding effects arising from the differing identity of the colonisers. I also do not include observations for the years 1775-1776, as the American Revolution was under way in this period. I furthermore exclude all the observations for Georgia and for Virginia before 1619. Since Georgia was established much later than the other colonies (1732) and slavery was forbidden until 1749, there were likely different processes that determined the evolution of representative institutions in this colony. In addition, Virginia was the first established colony, and there was a lag between settlement and the establishment of an assembly. Section 7 presents regression results which include all excluded years and observations.

I use the percentage of black population to capture the type of labour market (*Labour*) in a colony. Colonies that depended on white labour, either because there was a labour shortage as in the pre-slavery South or an adequate white labour supply as in the North, should have low values of this variable, while colonies that had a labour surplus should have a large relative share of the non-white population. There are three important critiques to this approach. First, Labour includes women, children and older individuals, and thus is an imperfect proxy for the black population actually in the labour force. Unfortunately, there are no available data for only working slaves or slaves above a certain age. Second, Labour lumps native-born and imported blacks. Finally, Labour may serve as a proxy for other variables (such as inequality) in addition to or instead of capturing labour markets. I address this and other validity concerns in Section 7.

To measure *Inequality*, I use the percentage of a colony's white landless population. Given that land was the major source of wealth in the colonial period and all colonies provided at least a subsistence-level quality of life to their white population, a large percentage landless should indicate that colonial resources were concentrated in the hands of a small elite, and vice versa.[26] Note that changes in land inequality can arise for two reasons: a change in the concentration among the landed, and changes in the concentration among the landless. The adopted strategy would be valid as long as the concentration among the landed stayed roughly the same across the colonies, and the only alteration of the overall distribution arose because of movement in the percentage landless category. In addition, percentage landless is unlikely to be a proxy for local labour markets, as landless farmers tended to rent land (in the South, they were often able to also buy slaves or indentured servants). As a result, free labour comprised a small percentage of labour supply in both the South and the North (McCusker and Menard 1991).

The proxy for Inequality is similar to the one used by Goldin and Katz (1999). The authors argue that, for the period 1910-1940, automobile registrations per capita summarise the distribution of wealth. In particular, in two states with the same mean wealth, the state with the more unequal distribution will have a greater proportion of households with automobiles. Given that in the period under consideration by the authors automobiles were luxury goods, such an interpretation

---

[25]Results are unchanged if one uses a three-year or a seven-year panel. I focus on these panels instead of calculating 5 or 10-year averages because the dependent variable is less persistent in the former specification. In the robustness section, I show that results are not sensitive to the type of panel chosen. The same approach is also adopted by Acemoglu et al (2008b) on the grounds that averaging introduces additional serial correlation, but in this analysis serial correlation is not an issue in either panel.

[26]As Risch (1937) explains, land was one of the greatest attractions offered in the New World and wealth, social position and political privilege were all related to it.

is logical. Since land in colonial America was widely available, it makes sense to assume that increases in percentage landless led to an increase in Inequality, and vice versa.[27]

As a measure of GDP per capita (*Income*), I use historical building wage data. The basic assumption behind using wages as a proxy for Income is that if wage is equal to marginal product, compensation for comparable occupations should be higher in places that are richer. As Margo (1992, 174-175) explains, subject to a few caveats (for example, the real wage index is broadly defined and wages actually represent the true budget constraint of the population), using a wage index to measure changes in the standard of living is appropriate.

Appendix 1 provides additional information on the data used in the regressions.

5 Empirical setup

5.1 Base model

The following econometric model summarises the relationship between suffrage and labour markets:

$$Suffrage_{ct} = \alpha + \beta_1 Suffrage_{ct-1} + \theta_1 Labour_{ct-1} + \theta_2 Inequality_{ct-1} +$$
$$+ \beta_2 Income_{ct-1} + Year_t + Latitude_c + \varepsilon_{ct} \tag{1}$$

where $Suffrage_{ct}$ is the suffrage index obtained from McKinley (1905) in colony c and time period t; $Suffrage_{ct-1}$ is *Suffrage Lagged One Period*; $Labour_{ct-1}$ is percentage black lagged one period; $Inequality_{ct-1}$ is percentage landless lagged one period; $Income_{ct-1}$ is wages lagged one period; $Year_t$ is a continuous year control; $Latitude_c$ is latitude; ε_{ct} is the disturbance term and α is the intercept.[28] The data set uses an unbalanced panel and covers the years 1619-1774. All independent variables are lagged by one period.[29]

Following the predictions of the theoretical model in Section 3 and the Supplementary Materials, a potential concern is reverse causality between Labour and Suffrage. While Labour has a direct effect on Suffrage, a change in Suffrage may subsequently affect Labour as well. There are two reasons why reverse causality will affect the magnitude of the Labour coefficient only marginally. First, the overall effect of Suffrage on Labour will be muted if a large proportion of white labourers have high re-entry or moving costs, as is the assumption in the formal model. Similarly, the impact of Suffrage on Labour will most likely happen with an individual-specific lag, and if such lags are un-correlated, any effect will be spread out across multiple time periods.

---

[27]The suggested historical relationship between land ownership and inequality in Maryland also confirms the validity of the percentage landless variable. Jordan (1987) argues that as inequality in Maryland began to grow around 1690-1700, there were more landless tenants.

[28]The inclusion of a lagged dependent variable is intended to capture persistency as well as mean reversion. Depending on the panel, one-period lags are either 5 or 10 years. I do not cluster the errors at the colony level in this section, as the number of clusters is only 13, but results are unchanged when the errors are clustered (see Section 7.2).

[29]The augmented Dickey-Fuller unit root test rejected the null of a unit root when I applied it separately to each colony time series.

Even so, I address concerns about reverse causality by instrumenting Labour in the following subsection.

Note also that I am unable to introduce region/colony fixed effects or region/colony LSDVs (least squares dummy variables) in any of the regressions because of the presence of a lagged dependent variable (see Beck and Katz 2004). Controlling for Latitude captures any persistent geographic and/or climate differences and should alleviate to a large extent the fixed effects problem. Still, I rerun the presented results by substituting Latitude with colony fixed effects, and, if anything, the results for Labour are stronger than those presented in Table 3. These additional results are available upon request. [30]

5.2 Instrumental Variables approach

In this subsection, I relax the assumption about the exogeneity of Inequality and Labour and instrument once lagged Inequality with *Inequality Lagged Two Periods*, and Labour with a crop suitability index (*CropIndex)*. There are many problems with the lagged inequality instrument, as both the long-run levels of inequality and the extent of the suffrage may be affected by common long-term factors, such as geography. Likewise, there may be reverse causality between long-run institutions and inequality (Rogowski and MacRae 2004). At the same time, coming up with a valid instrument for Inequality is a very difficult undertaking, and an instrument using lagged values at least partially mitigates any endogeneity and reverse causality concerns. Since I cannot resolve the issues with the Inequality instrument, the results when Inequality is instrumented should be treated as suggestive only.

When it comes to instrumenting Labour, I can avoid the pitfalls of a lagged-variable instrument and instead rely on a richer IV strategy. I construct a Crop Index capturing the propensity of each colony to produce one of the three main colonial export crops (tobacco, rice and wheat) multiplied by the respective colonial crop price. Although I do not have data on the *British* price of these three commodities during the time period in question, historians point out that the colonies were largely price takers (Purvis 1999, Kulikoff 1986), so it is unlikely that the prices used in the construction of the IV were influenced by factors specific to the colonies. The logic behind this approach is that once the prices for labour-intensive crops increase (or once these crops are discovered, as in the case of rice), landowners will have an incentive to look for an additional labour supply. If the supply of British indentured servants is largely inelastic (as was the case in the late 17th and early 18th centuries), farmers will turn to slaves, leading to an increase in the overall availability of labour. In addition, the adopted IV strategy can eliminate the bias arising from measurement error if the measurement errors of the instrument and the instrumented variable are orthogonal.

More precisely, I construct the instrument for Labour in the following way. Each colony's propensity to produce tobacco, rice and wheat is based on a combination of geographic characteristics, such as temperature, rainfall and soil.[31] In the absence of such data for colonial America (or

---

[30]The econometric solution to the fixed effects issue is to use the Arellano-Bond GMM estimator, which would first-difference equation (1) and then use lagged levels of the variable of interest as instruments. However, such an approach is very taxing on the data and it may also lead to weak instruments.

[31]See Harlan (1992) for the crop compatibility of different geographical zones, and Easterly (2007) for an

for the United States), I construct the time-invariant crop suitability weights for each colony by using the total harvested acres for each of the three crops in the period 1866-2007. Since cotton production superseded rice production in the Southern colonies in the late 18th century, I count cotton production as part of rice production. After creating the crop suitability weight for each colony, I multiply the weight for each crop by its respective price series, obtained from the US Census.[32]

To ensure comparability across colonies, I convert the crop prices to British pence per pound using exchange rates available from the US Census.[33] Although the broad agricultural products grown in a colony should be a function of its geographic endowments, the timing of crop growing and the narrower crop specialisation are likely to be exogenous. For instance, commercial rice production in South Carolina did not start until approximately 1696, although the colony was settled nearly 35 years earlier. Gray (1933) suggests that rice became popular in this colony only after a ship from Madagascar, which reached the colony by accident, brought a strain of rice with grains that were larger and suitable for growth in both wet and dry lands (278).

Because of the issues with the Inequality instrument, in the specifications below I do not combine the instruments in a single equation, which would be the correct way of doing double instrumentation with two endogenous variables. Since the results for Inequality should be treated as suggestive only, I prefer to present separate results when Labour is instrumented. The results using the double instrumentation approach are similar to the ones presented, albeit less precise, and are available upon request.

Therefore, the corresponding first stages in the IV regressions are:

$$Labour_{ct-1} = \alpha^F + \beta_1^F Suffrage_{ct-1} + \beta_2^F Inequality_{ct-1} + \beta_3^F Income_{ct-1} +$$
$$+ \beta_4^F CropIndex_{ct-1} + Year_t^F + Latitude_c^F + \varepsilon_{ct}^F \tag{2}$$

and

$$Inequality_{ct-1} = \alpha^F + \beta_1^F Suffrage_{ct-1} + \beta_2^F Income_{ct-1} + \beta_3^F Labour_{ct-1} +$$
$$+ \beta_4^F Inequality_{ct-2} + Year_t^F + Latitude_c^F + \varepsilon_{ct}^F \tag{3}$$

All variables are as defined above, except for $CropIndex_{ct-1}$, which represents the Crop Index lagged one period, and $Inequality_{ct-2}$, which is Inequality Lagged Two Periods. The identifica-

---

example of a similar type of crop index on the country level.

[32] An alternative to using an index would be to instead include dummies for each crop interacted with price (see Easterly and Levine 2003). However, such an approach is less precise and also produces a high correlation between the tobacco/wheat dummies. I also constructed an alternative index with crop weights based on the qualitative evidence presented in McCusker and Menard (1991), and it is reassuring that the correlation between the two indices is 0.7886.

[33] I do not separately deflate these price series by the available US-wide CPI index, since there is evidence that inflation rates varied by colony but there are no data on colony-specific CPI indices (Historical Statistics of the US, Millennial Edition). However, this should not be a big concern, because the correlation between the deflated and non-deflated crop indices is very high (0.9653 in the whole sample), and the IV results are therefore very similar. These results are available upon request.

tion restrictions are that:

$$Cov(CropIndex_{ct-1}, \varepsilon_{ct} | Suffrage_{ct-1}, Income_{ct-1}, Inequality_{ct-1}, Controls) = 0 \qquad (4)$$

and

$$Cov(Inequality_{ct-2}, \varepsilon_{ct} | Suffrage_{ct-1}, Income_{ct-1}, Labour_{ct-1}, Controls) = 0, \qquad (5)$$

where ε_{ct} is the residual error term in the second-stage regression, (1). This empirical strategy is valid as long as the left-hand-side variables do not have an effect on Suffrage that works directly, through some other variable in the model or through omitted factors.

There are several concerns arising from the choice of an IV for Labour. First, the identification restrictions may not be satisfied, as the instrument may affect the dependent variable through other channels. For instance, the Crop Index variable may be correlated with colonial revenues (arguably proxied by income per capita), which may have an independent effect on the extension of the suffrage. For example, tobacco and rice producing colonies may be richer and better equipped to afford the redistribution associated with democratic institutions, relative to poor colonies. However, not only is Income largely insignificant in the regressions, but Section 7.4 also shows, using tax data, that colonial revenues are a poor determinant of the right to vote. In addition, the correlation between Crop Index and Income is only 0.07 in the sample.

Similarly, the Crop Index may simply be a proxy for geography or climate. Although the correlations of the instrument with Latitude, Average Temperature and the standard deviation of temperature are high (-0.609, 0.557 and -0.410, respectively), they are not exact.[34] In addition, the Crop Index regressions already take into account any independent effect of geography by controlling for average temperature. Finally, in line with the Engerman-Sokoloff hypothesis, Crop Index may be a direct determinant of Inequality. This is unlikely to be the case for several reasons. First, the Crop Index measures colonial crop prices as well as crop suitability, and the former were influenced by exogenous market conditions in Britain and Europe in general. Second, the unconditional correlation between Crop Index and Inequality is merely 0.156 in the sample. Likewise, any impact of factor endowments (as captured by the Crop Index) implies that inequality should be highly persistent, a fact refuted by Chart 6, which shows that Northern and Southern inequality varied considerably during the period studied in this paper. I address these and other identification concerns more thoroughly in Section 7.

In addition, there are issues with the construction of the instrument itself. Ideally, I would prefer to use a measure of net crop profit instead of simply the crop price, but unfortunately this more detailed information is unavailable. In the period under consideration, the price of wheat followed a downward trend, while that for rice went upwards. However, the price for tobacco exhibits a U-shaped pattern with the minimum happening around 1700. This trend is consistent with evidence

---

[34]Including geographic covariates in the first-stage regressions is a conservative approach which biases the coefficient on Labour in the second-stage regression downward. As stated in Section 5.1, I am unable to include colony fixed effects because of the presence of a lagged dependent variable in the second-stage regression. Since the crop suitability of a colony is largely determined by geography anyway, an argument against including geographic controls/fixed effects in this case is that looking at the overall, rather than within-colony, variation is more meaningful.

presented in Kulikoff (1986), who argues that lower prices for the crop in the late 1600s made tobacco farmers substitute slaves for indentured servants in order to cut costs (see Charts S3A-C in the Supplementary Materials).

Despite these shortcomings, the theoretical discussion and the results presented in the next section show that that Crop Index is a valid instrument for Labour. Not only are the two variables strongly correlated in the first stage, as indicated by the relevant t-statistics and Kleibergen-Paap statistics, but the correlation between Crop Index and the regression residuals is less than 0.05, suggesting that the instrument does not belong in the main regression.

6   Quantitative results

I present descriptive statistics from the 5-year panel on the regional level, before and after 1700 in Table 1, and a snapshot of the rise of the black population in the South in Table 2. I next turn to regression analysis. Table 3 shows that across aggregations of the dependent variable and regression and panel specifications, the coefficients on Labour and Inequality remain largely unchanged and significant, with Labour having a stronger and more consistent impact. The effect of Inequality is even weaker in the robustness check in Table 4. Moreover, Labour has a strong quantitative effect: a 10 percentage point increase in this independent variable leads to a short-term increase in Suffrage of approximately 0.232, or close to 8.1 per cent. The cumulative effect is much stronger: a 10 percentage point increase in Labour leads to an increase in Suffrage of roughly 0.559 restrictions, or 19.6 per cent.[35]

In column (1) of Table 3, I regress Suffrage only on lagged Labour. The coefficient on the main variable of interest preserves its significance when I add additional variables, such as Income and Inequality in column (2). Columns (3) - (9) conduct various robustness checks. In column (3) and (4), I drop Labour from the regression to see if the coefficient on the Inequality variable changes significantly (column (3) only considers Inequality, while column (4) tests whether Inequality has a non-linear effect on Suffrage by including a term *Squared Inequality*). While the magnitude of the Inequality coefficient barely increases in columns (3) and (4), the Squared Inequality term is insignificant. Column (5) presents the results from instrumenting Labour, while in column (6) I instrument once lagged Inequality with Inequality Lagged Two Periods. In both specifications, the coefficients on Labour are significant, although in column (5) it is less precisely estimated.

Finally, in columns (7)-(9), I restrict the sample to only the South, only the North and only to years after 1700 (the latter sample includes better quality data relative to the overall sample but less variation to estimate the effect). The effect of Labour is larger in the South (column (7)) as compared with the North (column (8)). This is expected, since Labour exhibited much more variation both within and across the Southern colonies during the period in focus, relative to the North. While Inequality is significant at the 5 per cent level in four out of the eight specifications, its significance drops to the 10 per cent and 20 per cent level in columns (6)-(9), thus questioning the robustness of the variable. Another interesting result from the table is that richer colonies had

---

[35]These caclulations are based on the results in column 2 in Table 3. As a comparison, the same increase in inequality leads to a short-term increase in Suffrage of 5 per cent and to a long-term increase of 12.4 per cent, but Inequality loses significance in some specifications. The cumulative effect is calculated by dividing the coefficient on Labour by the quantity (1-coefficient on Suffrage Lagged One Period).

less democratic institutions, contrary to findings in the democracy-development literature. This finding is likely driven by the growth in Southern income per capita after the arrival of slavery at the beginning of the 18th century. Whereas wages improved due to the rise in the production and prices of export crops, the quality of representative institutions worsened in response to the labour supply shock.[36]

Next, I rerun the above regressions using Suffrage Dinkin as the dependent variable (Table 4). The results largely confirm the robustness of the findings in Table 3, with Inequality significant at the 5 per cent level in only one specification. I conduct additional robustness checks in the following section.

7 Robustness checks

In this section, I examine the potential endogeneity of labour markets and the role of inequality and the differential levels of debt across colonies. I also run supplementary regressions.

7.1 Was Southern slavery an exogenous event?

The correlation of Labour with omitted variables or colony-specific characteristics is a serious identification concern. Although I do address this issue by instrumenting Labour in the empirical section, the instrument is subject to the several criticisms discussed above. In this section, I provide empirical evidence that the arrival of slavery in the South was influenced by exogenous factors. In addition, I show that the regression results are not sensitive to the use of different measures of labour markets.

Three important events affected the spread of Southern slavery: the rising British share in the slave trade; the low profitability of West Indies sugar production; and the decrease in the supply of British indentured servants. While Britain participated in the slave trade as early as 1662, the royal monopoly overseeing this activity was badly managed, and private adventurers and the Dutch supplied the bulk of slaves in the early South.[37] Britain's influence in the slave trade increased in the late 17th century due to rising British supremacy in the international arena. In addition, in 1698 the British Parliament opened the royal slave trade to private merchants, provided that they paid certain benefits to the Royal African Company. Such developments made slaves more readily available in the 13 colonies (Gray 1933, 352-354).[38]

Next, I study whether declining sugar profitability in the Caribbean corresponded with the arrival of slavery in the Southern colonies. The uniformly high ratio of slave prices over sugar prices in the period 1662-1689 must have squeezed out planter profits in the Caribbean and, especially towards the end of the period, made planters less eager to acquire new slaves, thus increasing the

---

[36]Confirming these findings, research has in fact shown that in certain cases higher income per capita may lead to *more* restrictive institutions, especially in cases of high primary commodity prices. See for instance Bobonis (2008) and Acemoglu (2006).

[37]The only exception to this pattern was South Carolina, whose initial settlers came from Barbados and had access to some slaves before the British slave trade expanded.

[38]Gray (1933, 349) argues that the arrival of the Restoration in Britain also motivated the British government to discourage the servant trade in the period 1664-1684.

supply of slaves available to other plantation economies, such as the colonial South (Chart 3). Moreover, this increase in the availability of slaves in the mainland colonies was accompanied by a decrease in the supply of British indentured servants. As Chart 4 shows, in the period 1655-1705 average earnings in Britain increased, suggesting that potential immigrants to the colonies had a greater incentive to stay at home. In sum, exogenous events occurring in the late 17th century provided incentives for Southern planters to switch from indentured servants to slaves.

I then rerun the regressions in Table 3 using two alternative measures of Labour. First, I use Caribbean slave prices scaled by the Crop Index. Since the Caribbean was the largest slave market throughout the colonial period, the region set slave prices for the rest of the New World (see Eltis, Lewis and Richardson 2005, 677). I prefer to use Caribbean slave prices instead of their North American counterpart, as the latter are likely endogenous to domestic demand. Moreover, I divide Caribbean slave prices by the Crop Index, rather than by Caribbean sugar prices as in Chart 3, because I am now interested in explaining the fluctuations in the state of labour markets rather than simply when the labour shock occurred. Therefore, a high ratio of Caribbean slave prices over the Crop Index indicates a "tight" labour market, as slaves are very expensive relative to the income obtained from export crops. Conversely, a low ratio means that the labour market is relatively "easy", as the relative price of slaves is low and landowners can easily satisfy all (or most) of their labour demand. Chart S4 shows that this ratio closely tracks the movement of percentage black: it is low before 1700 and high after 1700.

Unfortunately, because Caribbean slave prices are not colony-specific, I am unable to use the detailed colony-level data set employed in the regressions in Section 6 and can only run the analyses on the regional level, which clearly limits the variation from which the coefficients are estimated and introduces measurement error. Since white workers met labour demand in the North throughout the colonial period, I assume that the price of slaves there was 0. Because of these shortcomings of the data, the number of observations drops significantly when I take every fifth observation, leading to imprecisely estimated coefficients. As a result, in addition to the specification of Table 3, I use an annual panel with 5-year and 3-year lags as well as a panel obtained by taking every third observation. Although the results are marginally significant (likely due to measurement error and to the fact that labour markets affect representative institutions with a lag longer than three years), they are always correctly signed and broadly confirm the predictions of Table 3 (Table S2).

Next, I keep the colony-level setup, but I instead employ a binary measure of labour markets, which is 0 in the South before 1700 and in the North, and 1 in the South after 1700. Although this is a less precise proxy, it aims to capture the impact of all three exogenous events affecting the supply of slaves in the South discussed above. Again, the results in Table S3 are consistent with those presented earlier. While there is evidence that both inequality and labour markets experienced an upward surge in the South after 1700, labour markets had an independent and significant impact on the rise of Southern suffrage (Table S3).[39]

In short, all three different measures of labour markets show that the variable is a robust predictor of Suffrage. Moreover, the two additional measures do not seem to be simply proxies for

---

[39]The results are very similar when the regional level panel from Table S2 is used in this specification. These results are available upon request.

inequality. The unconditional correlation between the ratio of Slave Prices and the Crop Index and Inequality, is only 0.31 in the sample and less than 0.03 for the South, while the respective numbers for the binary indicator measure are 0.44 and 0.02. As expected, the binary indicator measure is indeed much more highly correlated with Inequality, but that should not be a concern as Inequality is controlled for in all regressions.

7.2 Additional regression specifications

To further test the robustness of the results in Tables 3 and 4, in this subsection I experiment with different regression setups and include several additional control variables. None of these specifications affected the significance of the variables of interest (Labour and Inequality), but in a few instances the magnitude of both variables increased. In addition, the magnitude of the coefficients decreased significantly when I kept all years and observations, but such a result is expected because representative institutions are unlikely to be affected to a great extent by short-term changes in the independent variables. All of the results are available in the Supplementary Materials (Tables S4-S14).

First, I rerun the regressions in Table 3 with two additional panel setups: using a 10-year panel (Table S4) and a panel calculated by taking 5-year averages (Table S5). Moreovoer, I experiment with another aggregation of Suffrage calculated through principal component analysis (Table S6). The coefficients on Labour and Inequality are somewhat higher in the 10-year panel, indicating that both variables have a stronger long-term effect. In the 5-year average panel, the coefficients on Labour and Inequality are slightly smaller in magnitude, but this is likely because the dependent variable is more persistent. In the 10-year panel and in the principal components panel, the relationship between Crop Index and Labour is not as strong in the first stage, possibly due to the introduction of additional noise in these variables.

Furthermore, I experiment with different setups for the standard errors. Table S7 presents a Prais-Winsten regression with robust standard errors, and shows that the results are robust when the errors are assumed to follow an AR(1) process and to be heteroskedastic.[40] Still, both types of regressions assume that the dependent variable is measured continuously, which does not hold as Suffrage is non-negative and censored above by 6.5. The solution to this problem is a corner solution regression, which is empirically equivalent to a Tobit type I regression presented in Table S8 (see Wooldridge 2002). Again, there are no significant changes in the size or significance of the main coefficients.

Next, I include alternative measures of Income in the regressions. Exports per capita (*Exports*) may be a better measure of colonial well-being if the wage index (Income) is measured with error (Table S9).[41] Colonies were established for trading with Britain, and their development depended largely on the amount of primary goods exported to the mother country (Purvis 1999). At

---

[40]Other setups intended to account for spatial and temporal autocorrelation of the errors, such as clustered standard errors, Newey-West standard errors (with one, two or three lags), panel corrected standard errors (PCSEs) following Beck and Katz (2001), and the inclusion of an additional lag of the dependent variable, produced very similar results. In fact, while Labour remains significant and consistent in magnitude, Inequality loses significance in several of these specifications. These results are available upon request.

[41]I only rerun some of the regressions with Exports because the sample size decreases significantly.

the same time, Exports is at best an imperfect measure of how rich (or developed) a colony was, as trade only accounted for 9-12 per cent of GDP in the 13 colonies (Sheridan 1984, 43). Following Acemoglu et al (2002), I also use *Population Density* and *Urbanisation* as additional proxies for Income (Table S10).[42] Once Exports are included, the Labour coefficient drops slightly in magnitude but remains significant across specifications, while the size of the Inequality coefficient decreases more than five times and becomes insignificant. Similar to the results with Income, more urbanised colonies have *worse* representative institutions, although Urbanisation is significant in only four of the nine specifications, while population density is insignificant.

In addition, I incorporate a number of supplementary control variables in the regressions. In Table S11, I rerun the regressions from Table 3 with dummy variables for when a particular colony was proprietary or charter (*Dummies for Proprietary or Charter Colony*; royal colony is the omitted category). The logic is that colonies seeking to maximise revenue (such as proprietary or charter colonies) may have a stronger incentive to liberalise the Suffrage in order to attract migrants. As expected, the dummies are positive and often significant, but their inclusion actually increases the size of the Labour coefficient by approximately 0.5 restrictions, while rendering the Inequality coefficient the same.

I also collected colony-level data on the ancestral origins of the white population in 1790 and on the distribution of different types of churches in 1750, from which I calculated Herfindahl-Hirschman Index-type indices of *Ethnic Fractionalisation* and *Religious Fractionalisation*. If colonies were highly divided on these dimensions, it could have been more difficult for actors to coordinate on the democratic outcome, and more autocratic institutions would have emerged (see Alesina et al 1999 and Easterly and Levine 1997). The correlation between these two measures is very high (0.87). Although the Southern colonies were highly fractionalised on both dimensions (with the exception of Virginia, where over 75 per cent of the churches were Anglican), so were the Middle colonies. I next introduced separately these two controls and reran the regressions from Section 6. The coefficients on Ethnic Fractionalisation and Religious Fractionalisation are mostly insignificant and unstable across regression aggregations, without affecting the significance and magnitude of the coefficients on the rest of the variables (Tables S12 and S13).

Finally, in Table S14 I repeat the regressions from Table 3 using an annual panel and all available observations. Despite the more than 5-fold drop in magnitude and the imprecise estimation of some of the coefficients, the coefficients on Labour and Inequality are generally significant, showing that representative institutions start adjusting almost instantaneously to economic changes, albeit to a very small extent. Again, the significance of the Inequality coefficients is much weaker relative to those for Labour.

One critique of the regression approach is that it is dependent on a particular functional form and may miss important non-linearities in the relationship among Labour, Inequality and Suffrage. Moreover, the research setup of the paper implicitly assumes that slavery was a combination of two quasi-experiments (see the discussion in the Introduction). Although the slavery shock affected the Southern colonies around roughly the same time and did not have a significant impact

---

[42]These two measures are less defensible in the context of the 13 colonies, as the richest colonies (in the South) were the least urbanised and also less populated. The production of crops such as rice and tobacco made the establishment of large cities unnecessary (McCusker and Menard 1991).

on the Northern colonies, the South and the North differed in terms of Average Temperature, Income, Inequality and Suffrage before the arrival of slavery (Table 1). As a result, a comparison of Northern and Southern institutions after slavery may reflect the impact of pre-slavery differences.

The synthetic control group approach developed by Abadie et al (2003, 2007) provides a potential solution to these issues. This technique resembles statistical matching for observational studies and compares the political evolution of the South to a control group constructed from a weighted combination of Northern colonies, rather than all Northern colonies. Those colonies whose economic and political paths resemble most closely the pre-slavery path of the treatment group are weighted most heavily in the synthetic control group. Even though Table 2 shows that slavery arrived at different times in the Southern colonies, in order to maximise the number of observations, I aggregate all Southern colonies.[43] Chart 5 confirms that the rise in Southern Suffrage was driven by the post-slavery changes in Labour and not by the pre-1700 differences between the North and the South or the regression assumptions.

7.3 Inequality

This paper argues that in economies with tight labour markets, elites will provide liberal representative institutions in order to cater to potential labourers. In addition, in such situations the rich will tend to downplay the redistributional effects of high inequality. Once labour market pressures are weakened, considerations of the income distribution will weigh in and the suffrage may be tightened. Some of the proponents of an alternative view on the role of inequality include Engerman and Sokoloff (2000, 2002 and 2006), Boix (2003), Boix and Stokes (2003) and Acemoglu and Robinson (2005). If Labour is simply a proxy for Inequality in the analysis, the presented results will be spurious.

If the arrival of slavery led to an increase in the distribution of income because it gave the rich an opportunity to establish large plantations, then such a relationship should be particularly pronounced in the Southern colonies. Graphical evidence shows that there is a weak positive but insignificant relationship between Inequality and Labour in the South (the correlation for the whole region is 0.1146; see Charts S5 and S6 in the Supplementary Materials). However, looking at simple correlations may mask the true relationship between inequality and slavery if there is reverse causality or omitted variables affecting both Inequality and Labour, such as geography. Regressing (for the South only) Inequality on its lagged value, Labour and controls for Year, Income and Average Temperature also produced an insignificant relationship. Finally, very similar results were obtained when the graphs and regressions were expanded to include the Northern colonies as well (Chart S7). Although the unconditional correlation between Inequality and Labour in the full sample is somewhat higher than that in the South (0.3162), the regression results again show the lack of a relationship between the two variables.

The joint evolution of Inequality and Suffrage in the North further questions the importance of only the income distribution in determining representative institutions. While Northern suffrage restrictions remained roughly constant throughout the colonial period, Northern inequality was decreasing until around 1700, after which it started increasing (Chart 6, compare with Chart 1). In fact, Jones (1980) shows that New England had the greatest wealth discrepancy in pre-

---

[43]I experiment with three different treatment years: 1690, 1700 and 1710, yielding very similar results.

Revolutionary America: in 1774, the richest 10 per cent of people who left probate records held 57 per cent of the net worth in New England, compared with 49 per cent in the South and 42 per cent in the Middle colonies.

7.4 Differential levels of debt and costs of liberal institutions across colonies

If the decision to tighten Southern suffrage following the arrival of slavery was motivated by the elites' desire to avoid the redistribution costs of liberal representative institutions, the presented results will be biased. Using data on colonial taxation compiled by Rabushka (2008) for the years 1688, 1714 and 1739, I argue that the theoretical and empirical findings of the paper are robust to this mechanism. For the sake of space, I only present a verbal discussion of the data.

The level of taxation in the 13 colonies was much lower than that in Britain, and therefore the colonies were largely debt-free. Although London prohibited the issuance of colonial paper money, the colonies managed to issue quasi- official money and to keep taxes low. New England developed relatively early a well-organised system of local taxation, including property and poll taxes, and along with some of the Middle colonies (such as New York), was the most heavily taxed region throughout the colonial period. Colonial governments in the Upper and Lower South raised their revenues mainly from export duties on tobacco and rice, and thus the level of taxation depended largely on the amount and value of export crops produced. As primary commodity prices increased in the 18th century, the Southern colonies could in fact afford *more* liberal institutions. I conclude that the concern about the costs of democratic institutions did not motivate the decision of Southern elites to restrict the suffrage following the arrival of slavery.

8 Conclusion

This paper studies why elites alter the suffrage using the case study of the 13 British American colonies. The advanced theory builds on the observation that the right to vote is an important incentive for workers not only because of its short-term monetary benefits, but also because it serves as a commitment device for any promises made by those in power. Therefore, when designing representative institutions elites will take into account both inequality and the state of labour markets. When there is a labour shortage, democratic institutions will emerge even in cases of high inequality because such a policy will attract additional workers. In contrast, under a labour surplus those in power will opt for more autocratic political institutions as concessions to potential migrants are no longer warranted.

The theoretical framework can easily be applied to other contexts and time periods. For example, female enfranchisement in Western Europe and the United States coincided with the end of the First World War, which made male workers scarce.[44] Similarly, countries with a labour shortage in particular occupations have point-based immigration schemes that grant citizenship and the associated political rights to qualified candidates. For instance, Canada and Australia have special programmes giving permanent residence to highly skilled immigrants.

---

[44]In the same vein, recent research argues that US states which were short on women were quicker in adopting universal suffrage (see Braun and Kvasnicka 2009). Using the evidence in Section 3.2, a low number of women, relative to men, can be interpreted as a proxy for tight labour markets.

As far as methodology is concerned, this paper adds to the studies that break away from the cross-country regression approach and focus on a single country or region in order to better determine causality. Moreover, I improve on such designs by uniquely tracing the economic and political evolution of the 13 colonies starting from the first years of their settlement to the American Revolution. In addition to putting forward a new mechanism linking labour markets and representative institutions, this work also provides a novel research setup to test the developed theory.

References

A. Abadie and J. Gardeazabal (2003), "The economic costs of conflict: a case study of the Basque country", *American Economic Review*, 93(1), pp. 113-132.

A. Abadie, A. Diamond and J. Hainmueller (2007), "Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program", NBER Technical Working Papers: 0335.

D. Acemoglu, 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(4), pp. 1231-1294.

D. Acemoglu and J. A. Robinson (2005), *Economic Origins of Dictatorship and Democracy*, Cambridge University Press.

D. Acemoglu (2006), "Modeling inefficient economic institutions", in R. Blundell, W. Newey and T. Persson (eds), *Advances in Economic Theory, Proceedings of 2005 World Congress*, Cambridge University Press.

D. Acemoglu, M. A. Bautista, P. Querubin and J. A. Robinson (2008a), "Economic and political inequality in development: the case of Cundinamarca, Colombia", in E. Helpman (ed.), *Institutions and Economic Performance*, Harvard University Press.

D. Acemoglu, S. Johnson, J. A. Robinson and P. Yared (2008b), "Income and democracy", *American Economic Review*, 98(3), pp. 808-42.

D. Albouy (2008), "The colonial origins of comparative development: an investigation of the settler mortality data", NBER Working Paper No. 14130.

A. Alesina, R. Baqir and W. Easterly (1999), "Public goods and ethnic divisions", *Quarterly Journal of Economics*, 114(4), pp. 1243-1284.

A. Atkinson and A. Brandolini (2001), "Promise and pitfalls in the use of "secondary" data sets: income inequality in OECD countries as a case study", *Journal of Economic Literature*, 39(3), pp. 771-799.

T. Ballard, Email communication. USDA National Agricultural Statistics Service Statistics Divisions, Crops Branch, Field Crops Section, April 13, 2009.

A. Banerjee and L. Iyer (2005), "History, institutions and economic performance: the legacy of colonial land tenure systems in India", *American Economic Review*, 95(4), pp. 1190-1213.

S. Basher and N. Lagerlöf (2008), "Per-capita income gaps across US states and Canadian provinces", *Journal of Macroeconomics*, 30(3), pp. 1173-1187.

N. Beck and J. Katz (2001), "Time-series cross-section data: what have we learned in the past few years?" *Annual Review of Political Science*, 4(2001). pp. 271-293.

T. Beck and L. Laeven (2006), "Institution building and growth in transition economies", *Journal of Economic Growth*, 11, pp. 157-186.

N. Beck and J. Katz (2004), "Time-series-cross-section issues: dynamics, 2004", working paper.

G. Bertocchi and C. Strozzi (2008), "International migration and the role of institutions", *Public Choice*, 137(1), pp. 81-102.

D. Bloom and J. Sachs (1998), "Geography, demography and economic growth in Africa", in William Brainard and George Perry (eds), *Brookings Papers on Economic Activity, Volume 2*, Brookings Institution Press.

G. Bobonis (2008), "Endowments, coercion and the historical containment of education", University of Toronto working paper.

C. Boix (2003), *Democracy and Redistribution*, Cambridge University Press.

C. Boix and S. Stokes (2003), "Endogenous democratization", *World Politics*, 55(4), pp. 517-49.

C. Boix (2010), "From income to democracy?", working paper.

S. Braun and M. Kvasnicka (2009), "Men, women, and the ballot - woman suffrage in the United States", Ruhr Economic Paper N 93.

R. Brown (1969), *Middle-Class Democracy and the Revolution in Massachusetts, 1691-1780*, Harper Torchbooks.

R. Dinkin (1977), *Voting in Provincial America: A Study of Elections in the Thirteen Colonies, 1689-1778*, Greenwood Press.

R. Dunn (2000), *Sugar and Slaves: The Rise of the Planter Class in the English West Indies, 1624-1713*, University of North Carolina Press.

W. Easterly and R. Levine (1997), "Africa's growth tragedy: policies and ethnic divisons", *Quarterly Journal of Economics*, 112(4), pp. 1203-1250.

W. Easterly and R. Levine (2003), "Tropics, germs and crops: how endowments influence economic development", *Journal of Monetary Economics*, 50(1), pp. 3-39.

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

L. Edlund (2005), "Sex and the city", *Scandinavian Journal of Economics*, 107(1), pp. 25-44.

S. Engerman and K. Sokoloff (2000), "History lessons: institutions, factor endowments, and paths of development in the New World", *Journal of Economic Perspectives*, 14, pp. 217-232.

S. Engerman and K. Sokoloff (2002), "Factor endowments, inequality, and paths of development among New World economies", *Economía*, 3(1), pp. 41-88.

S. Engerman and K. Sokoloff (2006), "Colonialism, inequality and long-run paths of development", in A. Banerjee, R.Benabou, and D. Mookherjee (eds.) *Understanding Poverty*, Oxford University Press.

R. W. Fogel (1989), *Without Consent or Contract: The Rise and Fall of American Slavery*, W.W. Norton and Company.

G. Friedman (1989), "The rise of the New South and the geographic regions of cotton, rice and sugar", in Fogel, Robert W., Ralph A. Galantine and Richard L. Manning, eds. *Without Consent and Contract: The Rise and Fall of American Slavery, Evidence and Methods*, W.W. Norton and Company, pp. 192-195.

D. Galenson (1984), "The rise and fall of indentured servitude in the Americas: an economic analysis", *Journal of Economic History*, 44(1), pp. 1-26.

D. Galenson (1996), "The settlement and growth of the colonies: population, labour and economic development" in Engerman, Stanley and Robert E. Gallman, eds. *The Cambridge Economic History of the United States. Volume 1*, pp. 135-208.

C. Goldin and L. F. Katz (1999), "The shaping of higher education: the formative years in the United States, 1890-1940", *Journal of Economic Perspectives*, 13(1), pp. 37-62.

L.C. Gray (1933), *History of Agriculture in the Southern United States To 1860, Volume I*. A.M.Kelley.

S. J. Grossman and O.Hart (1986), "The costs and benefits of ownership: a theory of vertical and lateral integration", *Journal of Political Economy*, 94(4), pp. 691-719.

O. Hart and J. Moore (2007), "Incomplete contracts and ownership: some new thoughts", *AEA Papers and Proceedings*, 97(2), 182-86.

J. Harlan (1992), *Crops and Man*, American Society of Agronomy-Crop Science Society.

S. Jha (2008), "Shareholding, coalition formation and political development: evidence from 17th century England", Stanford University, Graduate School of Business, Research Papers.

D. Jordan (1987), *Foundations of Representative Government in Maryland: 1632-1715*, Cambridge University Press.

A. H. Jones (1980), *Wealth of a Nation to Be*, Cambridge University Press.

A. Keyssar (2000), *The Right to Vote: The Contested History of Democracy in the United States*, Basic Books.

D. Klingaman (1969), "The significance of grain in the development of the tobacco colonies", *Journal of Economic History*, 29(2), 268-278.

A. Kulikoff, Alan (1986), *Tobacco and Slaves: The Development of Southern Cultures in the Chesapeake, 1680-1800*, University of North Carolina Press.

A. Kulikoff, 2000, *From British Peasants to Colonial American Farmers*, University of North Carolina Press.

S. M. Lipset (1959), "Some social requisites of democracy: economic development and political legitimacy", *American Political Science Review*, 53(1), 69-105.

A. Lizzeri and N. Persico (2004), "Why did the elites extend the suffrage? Democracy and the scope of government, with an application to Britain's 'age of reform,' " *Quarterly Journal of Economics*, 119(2).

R. Margo (1992), "Wages and prices during the antebellum period: a survey and new evidence", in R. Gallman and J. Wallis, eds. *American Economic Growth and Standards of Living before the Civil War*, Chicago: University of Chicago Press, pp. 173-210.

J. McCusker and R. Menard (1991), *The Economy of British America, 1607-1789*, University of North Carolina Press.

A. H. Meltzer and S. F. Richard (1981), "A rational theory of the size of government", *Journal of Political Economy*, 89(5), pp. 914-927.

E. Miguel (2004), "Tribe or nation? Nation-Building and public goods in Kenya versus Tanzania", *World Politics*, 56(3), pp. 327-362.

E. Morgan (1975), *American Slavery, American Freedom: The Ordeal of Colonial Virginia*, W.W. Norton and Company.

G. Nash (1979), *The Urban Crucible: Social Change, Political Consciousness, and the Origins of the American Revolution*, Harvard University Press.

N. Nunn, Nathan (2008), "Slavery, inequality, and economic development in the Americas: an examination of the Engerman-Sokoloff hypothesis", in E. Helpman (ed.), *Institutions and Economic Performance*, Harvard University Press, pp. 148-180.

R. Pande and C. Udry (2005), "Institutions and development: a view from below", in R. Blundell, W. Newey, and T. Persson (eds.), *Proceedings of the 9th Congress of the Econometric Society*, Cambridge University Press.

E. Perkins (1980), *The Economy of Colonial America*, Columbia University Press.

A. Przeworski (2007)", Conquered or granted? A history of suffrage extensions", NYU Department of Politics working paper.

T. Persson and G. Tabellini (2006), "Democracy and development: the devil in the details", *American Economic Review*, 96(2), pp. 319-324.

A. Przeworski and F. Limongi (1997), "Modernization: theories and facts", *World Politics*, 49(2), pp. 155-183.

T. Purvis (1999), *Colonial America to 1763*, Facts on File, Inc.

T. Purvis (1994), *Revolutionary America, 1763 to 1800*, Facts on File, Inc.

A. Rabushka (2008), *Taxation in Colonial America*, Princeton University Press.

R. Rogowski and D. MacRae (2004), "Inequality and institutions: what theory, history and (some) data tell us", Paper prepared for the 2004 APSA conference, September 2-4, Chicago.

G. Roland (2002), "The political economy of transition", *Journal of Economic Literature*, 16(1), pp. 29-50.

E. Risch (1937), "Encouragement of immigration: as revealed in colonial legislation", *The Virginia Magazine of History and Biography*, 45(1), pp. 1-10.

D. Ryden and R. Menard (2005), "South Carolina's colonial land market: an analysis of rural property sales, 1720-1775", *Social Science History*, pp. 29(4), pp. 599-623.

R. Sheridan (1984), "The domestic economy", in Jack Greene and J.R. Pole (eds.), *Colonial British America: Essays in the New History of the Early Modern Era*, Johns Hopkins University Press, pp. 43-85.

K. Sokoloff and G. Villaflor (1982). "Migration in colonial America: evidence from the militia muster rolls", *Social Science History*, 6(4), pp. 539-570.

D. Ticchi and A. Vindigni (2008), "War and endogenous democracy", IZA working paper 3397.

D. A. Williams (1969), "The Small farmer in eighteenth-century Virginia politics", *Agricultural History*, 43(1), pp. 91-102.

G. Wright (2006), *Slavery And American Economic Development*, Louisiana State University Press.

J. Wooldridge (2002), *Econometric Analysis of Cross-Section and Panel Data*, MIT Press.

Data Sources

Bureau of Labour Statistics (1929), *History of Wages in the United States from Colonial Times to 1928*, United States Government Printing Office.

Census Bureau Tiger Map, http://www.census.gov/geo/www/tiger/latlng.txt (accessed 10 October, 2011).

Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, http://www.noaa.gov/ (accessed 5 April, 2009).

Distance Calculator, http://www.infoplease.com/atlas/calculate-distance.html (accessed 27 October, 2008).

D. Eltis, F. D. Lewis and D. Richardson (2005), "Slave prices, the African slave trade, and productivity in the Caribbean, 1647-1807", *Economic History Review*, LVIII(4), pp. 673-700.

Historical Census Browser, Retrieved 1 August, 2008, from the University of Virginia, Geospatial and Statistical Data Center: http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html.

J. T. Main (1965), *The Social Structure of Revolutionary America*, Princeton University Press.

H. Moller (1945), "Sex composition and correlated culture patterns of colonial America", *The William and Mary Quarterly, Third Series*, 2(2), pp. 114-153.

National Agricultural Statistics Service, United States Department of Agriculture, Retrieved 1 August, 2008. http://www.nass.usda.gov/Data_and_Statistics/Quick_Stats/index.asp.

North American ports, Wikipedia, Retrieved 1 October, 2011.

Officer, Lawrence (2008), "The Annual RPI and Average Earnings for the United Kingdom, 1264 to 2007", MeasuringWorth.com, (accessed April 5, 2009).

T. Purvis (1994), *Revolutionary America, 1763 to 1800*, Facts on File, Inc.

R. Sutch and S. Carter (eds.) (2006), *Historical Statistics of the United States, Millennial Edition Online*, Part E, Government and International Relations, Colonial Statistics.

United States. Bureau of the Census (1975), *Historical Statistics of the United States, Colonial Times to 1970*, US Department of Commerce, Bureau of the Census, US Government Printing Office.

Chart 1: Evolution of the suffrage in the 13 colonies, by region
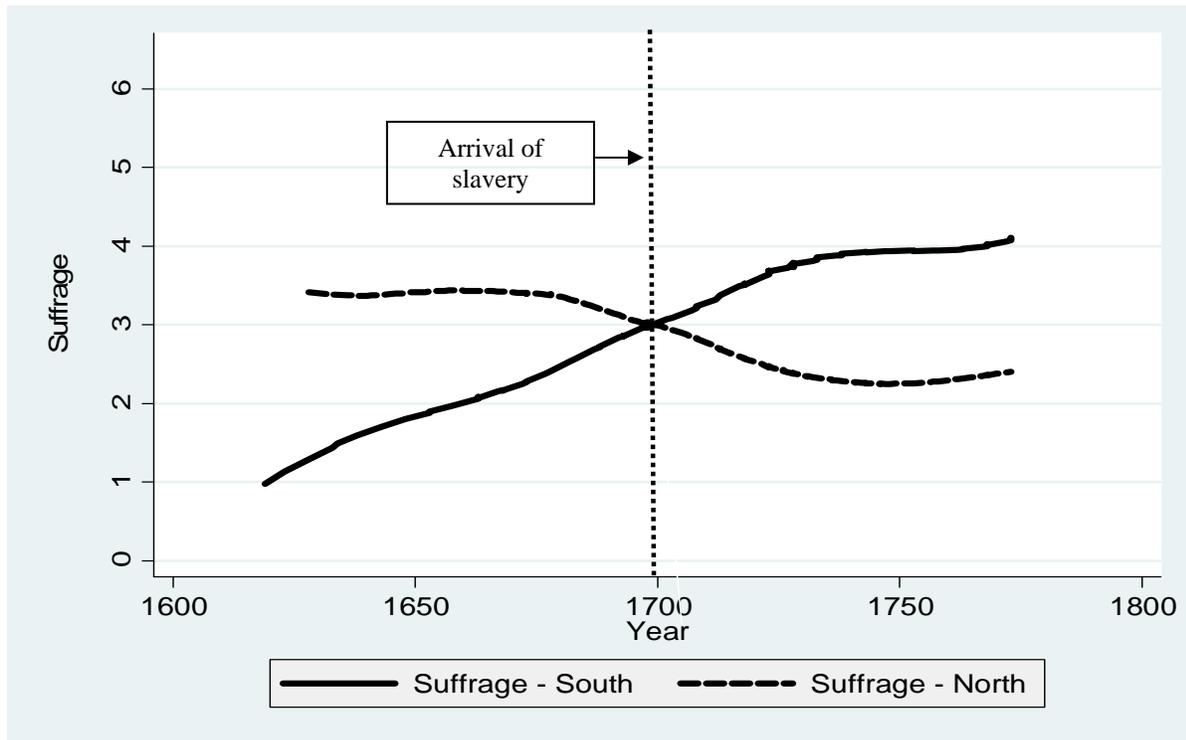
Source: Author's calculations from McKinley (1905).

Notes: This graph shows how suffrage restrictions in the North and South evolved over time. The category "South" includes Virginia, Maryland, South Carolina and North Carolina, while the category "North" includes Massachusetts, Rhode Island, Maine, Connecticut, New Hampshire, New York, New Jersey, Pennsylvania and Delaware. The lines are obtained by locally weighted least squares smoothing over all colony-year suffrage observations for a particular region (solid line = South, dashed line = North). Suffrage is calculated by summing up dummies for different suffrage restrictions, such as land ownership, income and religion. Suffrage ranges from 0 to 6.5, and a higher value of Suffrage indicates a more restrictive political system. Restrictions that cannot be substituted with another restriction are coded as 1; restrictions that can be substituted with one other restriction are coded as 1/4; restrictions that can be substituted with two other restrictions are coded as 1/9, and periods of no elections are coded as having the maximum number of suffrage restrictions in the sample. The data set is obtained by taking every fifth observation for each colony. See Appendix 1 and the text for more details on the construction of the variables.

Chart 2: Sex ratio and suffrage
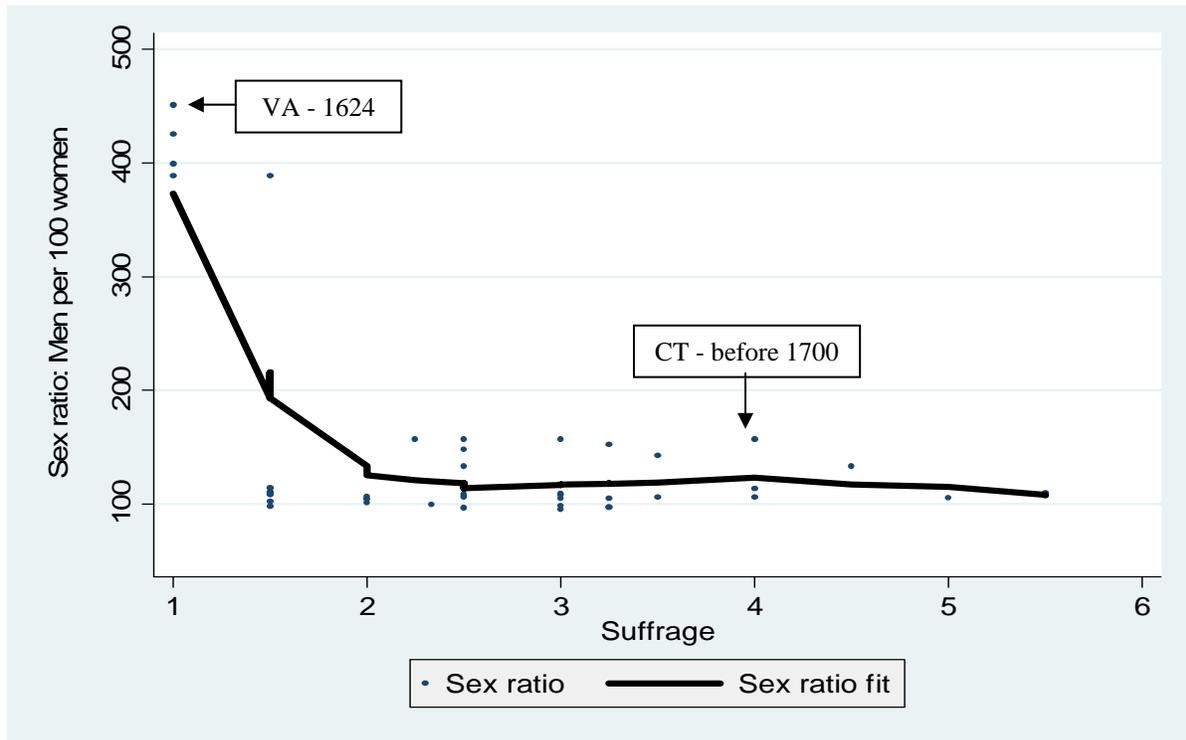
Sources: Sex ratio data: Wells (1975), Moller (1945); Suffrage data: McKinley (1905).

Notes: This graph shows the relationship between Suffrage and the Sex Ratio using variation over time and across colonies. The solid line is obtained by locally weighted least squares smoothing over all colony-year observations with data available for both variables. The sex ratio is calculated as the number of men per 100 women. If there was a discrepancy in the two sex ratio data sources, I used the data from Wells (1975). Because of the lack of data, sex ratios in 1775 were assumed to be the same as those in 1790. Missing data for this variable before 1700 were assigned based on region (for example, NC before 1700 was assigned the sex ratio value for SC for 1672). Sex ratio observations for the same colony that were less than 8-10 years apart were grouped together using arithmetic averaging. See the notes to Chart 1 and Appendix 1 on how Suffrage is constructed.

Chart 3: Ratio of Caribbean slave prices over sugar prices, 1638-1774
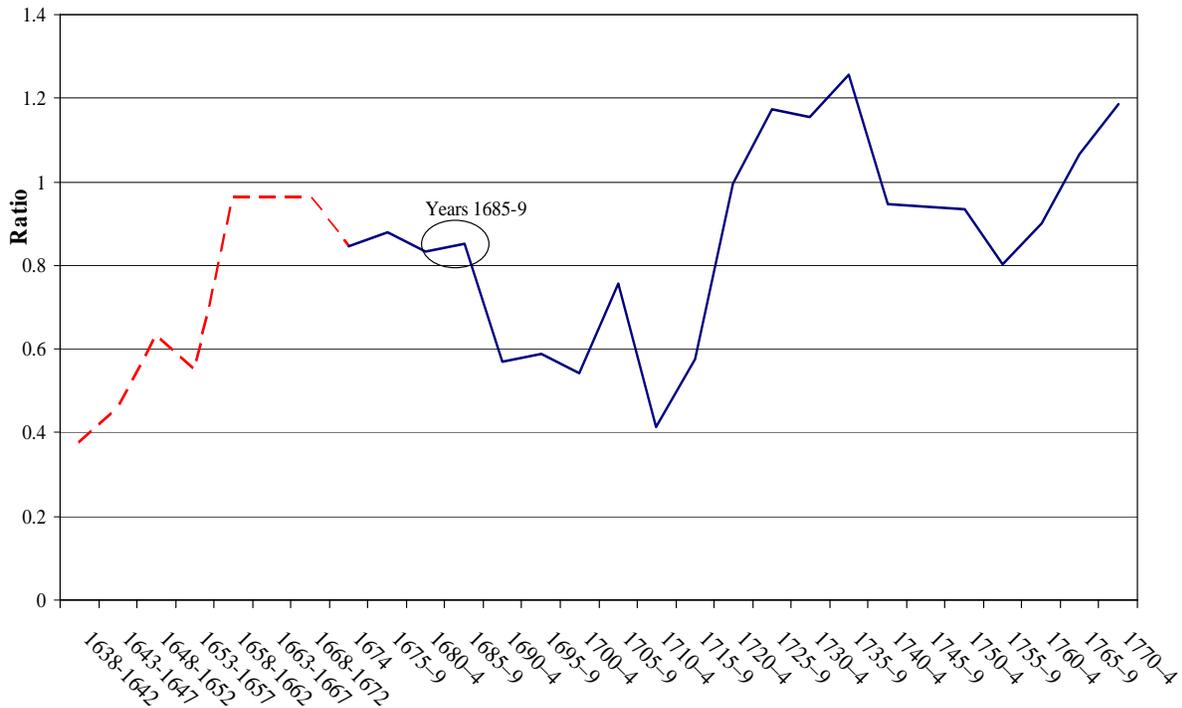
Source: Eltis, Lewis and Richardson (2005).

Notes: This graph shows how the ratio of Caribbean slave prices over sugar prices changed over time. Slave prices are in British pounds sterling, and sugar prices are in British pence per hundredweight (1 hundredweight = 112 lbs). The solid line (1674-1774) uses data on Caribbean slave and sugar prices from Eltis, Lewis and Richardson (2005), Table 2, 679. The dashed line uses British-American slave prices from Historical Statistics of the US (1975) and estimates Caribbean sugar prices based on the discussion in the Eltis et al article (677-678).

Chart 4**:** Average real earnings in the United Kingdom, 1600-1775
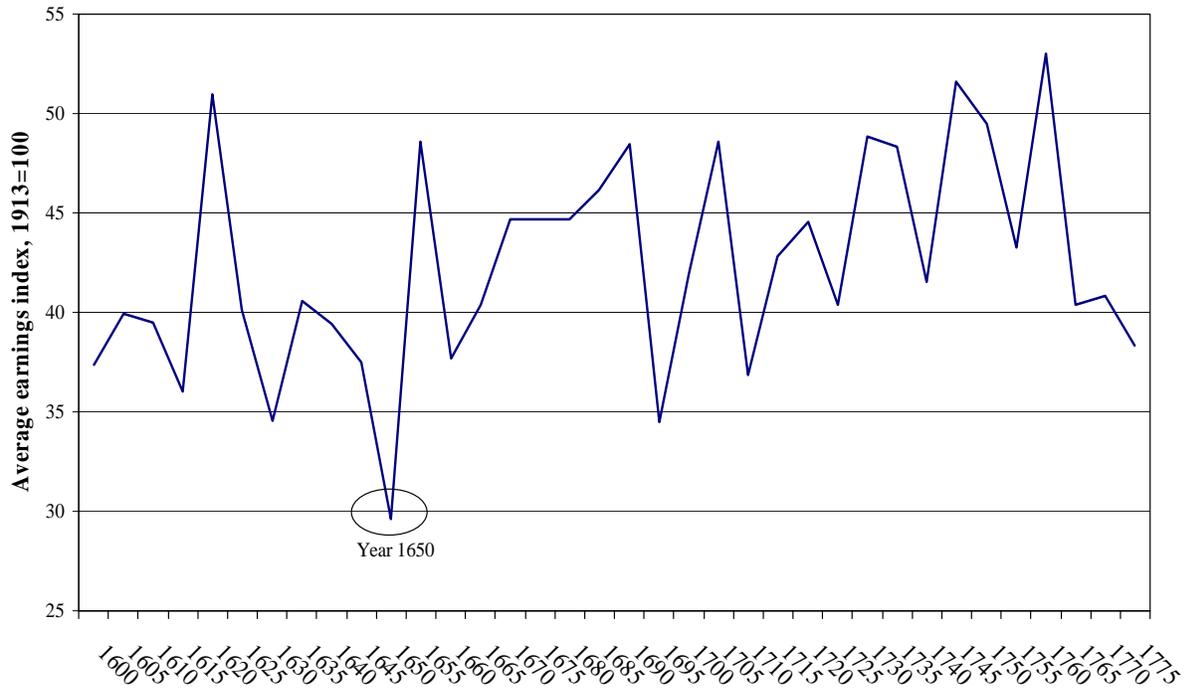
Source: Author's calculations from Officer, MeasuringWorth (2008), "The Annual RPI and Average Earnings for the United Kingdom, 1264 to 2007".

Notes: This graph shows the evolution of average real earnings in the United Kingdom in the 17th and 18th centuries. The chart uses five-year averages in order to minimise data volatility. The wage index is normalised so that 1913=100.

Chart 5: Evolution of suffrage in the South versus suffrage for a synthetic control group
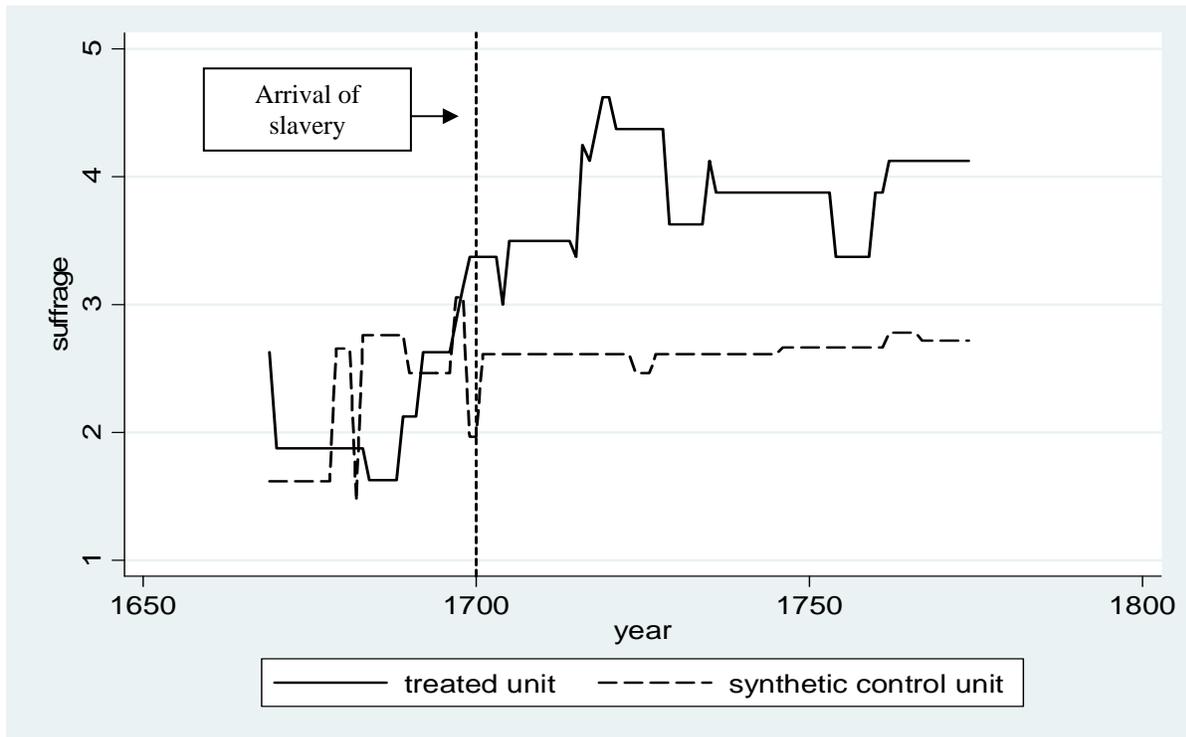
Source: Author's calculations from McKinley (1905).

Notes: This graph shows how the evolution of Suffrage in the South compares with that for a control group constructed from those Northern colonies whose economic and political characteristics most closely approximated the pre-slavery path of the South. The procedure constructs the synthetic control group based on pre-1700 similarity in the values of Suffrage averaged from 1670 to1680, from 1680 to 1690 and from 1690 to 1700; of Labour lagged one period; and of Inequality lagged one period. The treated unit is considered to be the whole Southern region, and the treatment period is assumed to be 1700. The synthetic control group consists of a weighted average of NH, NY and RI, where the weights are 0.297, 0.647 and 0.056 respectively. The root mean squared prediction error (RMSPE) is 0.663. The graph uses a balanced annual panel from 1669 onwards, and excludes Georgia, Maine and Pennsylvania, as these colonies do not have enough observations for the covered period.

Chart 6: Evolution of inequality – North and South
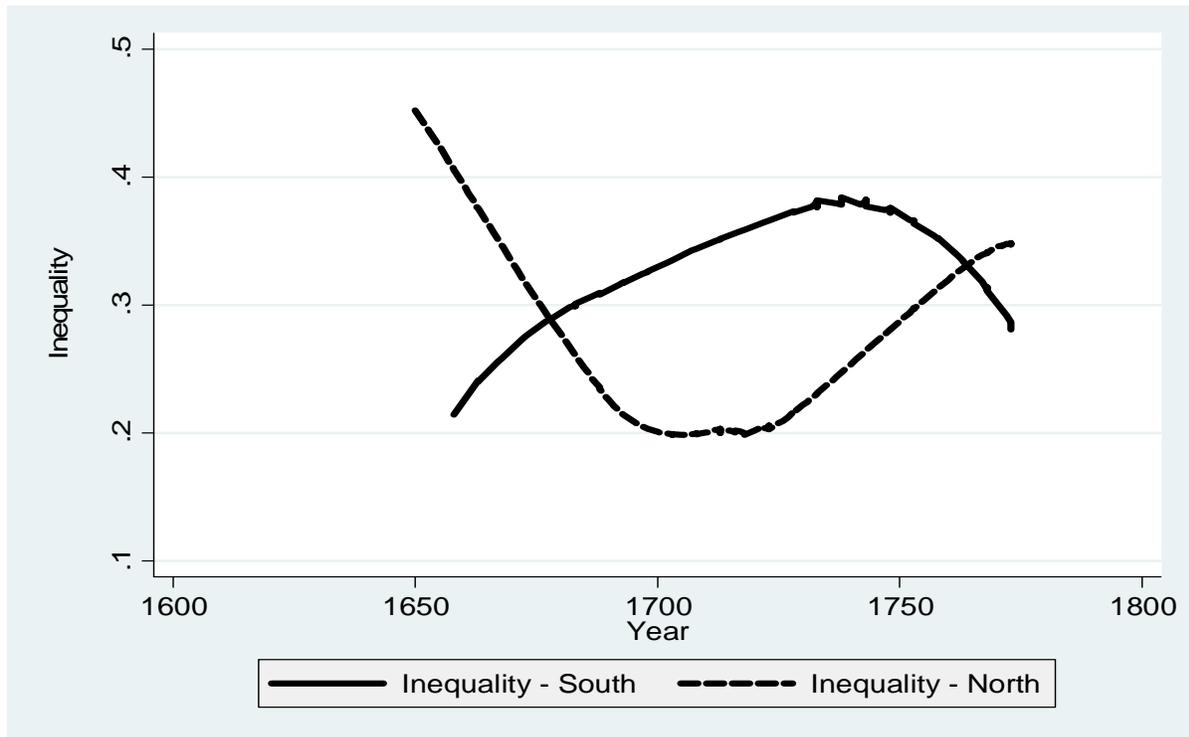
Sources: Author's calculations from Kulikoff (1986, 2000), Main (1965) and Nash (1979).

Notes: This graph shows how inequality in the North and South changed over time. See the notes to Chart 1 about the colonies included in each region. Inequality is measured as percentage landless, where a higher value of percentage landless implies higher inequality. The data set is obtained by taking every fifth observation for each colony. The lines are obtained by locally weighted least squares smoothing over all colony-year inequality observations for a particular region (solid line = South, dashed line = North). See Appendix 1 and the text for more details on the construction of the variables.

Table 1**:** Summary Statistics, five-year panel

	Before 1700					After 1700				
	North	**South**	**Difference**	**N North**	**N South**	**North**	**South**	**Difference**	**N North**	**N South**
	Panel A: Geographic characteristics									
Latitude	41.81	37.14	4.672***	86	46	41.43	36.59	4.841***	119	60
Average Temperature	48.10	56.57	-8.467***	86	46	48.76	57.67	-8.916***	119	60
	Panel B: Political and economic characteristics									
Suffrage	3.47	2.05	1.417***	86	46	2.34	3.89	-1.553***	119	60
Suffrage Dinkin	2.82	1.17	1.646***	86	46	2.28	3.34	-1.062***	119	60
Inequality	0.26	0.30	-0.037	46	27	0.25	0.36	-0.113***	119	60
Income	49.78	31.02	18.761***	78	32	69.82	63.16	6.658**	119	59
Labour	0.04	0.09	-0.059***	85	45	0.06	0.35	-0.297***	119	60
Crop Index	0.26	0.38	-0.126***	72	37	0.25	0.60	-0.344***	119	60
Exports	0.40	1.14	-0.747**	9	7	0.28	1.49	-1.211***	89	60
	Panel C: Demographic and other characteristics									
Ethnic Fractionalisaton	0.58	0.35	0.230***	86	46	0.51	0.33	0.173***	119	60
Religious Fractionalisation	0.54	0.43	0.104***	83	46	0.47	0.39	0.076**	119	60
Population Density	1.42	0.62	0.795***	85	45	10.08	4.63	5.451***	119	60
Urbanisation	0.15	0.04	0.106***	86	46	0.09	0.04	0.052***	119	60
Dummy for Proprietary Colony	0.43	0.59	-0.157*	86	46	0.25	0.37	-0.115	119	60
Dummy for Charter Colony	0.31	0.04	0.270***	86	46	0.24	0.00	0.244***	119	60
Dummy for Royal Colony	0.26	0.37	-0.114	86	46	0.50	0.63	-0.129	119	60

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows summary statistics for the main variables used in the regressions. The category "South" includes Virginia, Maryland, South Carolina and North Carolina, while the category "North" includes Massachusetts, Rhode Island, Maine, Connecticut, New Hampshire, New York, New Jersey, Pennsylvania and Delaware. The data set is obtained by taking every fifth observation for each colony. ***$p<=0.01$, **$p<=0.05$, *$p<=0.1$.

Table 2: Estimated percentage of black population in
the South, 1610-1780

	Upper South		Lower South		
Year	Virginia	Maryland	North Carolina	South Carolina	Georgia
1610	0.00%				
1620	0.91%				
1630	2.00%				
1640	1.44%	3.43%			
1650	2.16%	6.66%			
1660	3.52%	9.00%	2.00%		
1670	5.66%	9.00%	3.90%	15.00%	
1680	6.88%	9.00%	3.87%	16.67%	
1690	17.62%	9.00%	3.95%	38.46%	
1700	27.99%	10.90%	3.87%	47.92%	
1710	25.97%	18.59%	5.95%	55.34%	
1720	30.26%	18.90%	14.10%	64.54%	
1730	30.35%	18.90%	20.00%	66.67%	
1740	33.25%	20.70%	21.25%	72.32%	0.00%
1750	45.25%	30.80%	27.13%	66.22%	19.23%
1760	41.38%	30.20%	30.38%	60.95%	37.36%
1770	41.97%	31.50%	35.29%	60.51%	45.45%
1780	41.00%	33.47%	33.69%	53.89%	37.15%

Source: Historical Statistics of the US Millennial Edition Online (2006), Table Eg1-59.

Notes: This table shows the growth of the black population in the South in the late 17th
and early 18th centuries. Slavery was forbidden in Georgia until 1749.

Table 3: Suffrage in the 13 colonies: five-year panel, main specification

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.645*** [0.064]	0.586*** [0.093]	0.671*** [0.078]	0.683*** [0.075]	0.549*** [0.123]	0.717*** [0.069]	0.252 [0.168]	0.692*** [0.099]	0.648*** [0.092]
Labour	**2.041*** [0.587]**	**2.315*** [0.729]**			**3.178 [2.033]**	**1.368*** [0.521]**	**2.352** [0.979]**	**1.501* [0.882]**	**1.694*** [0.614]**
Income		0.004** [0.002]	0.005* [0.002]	0.004 [0.002]	0.004* [0.002]	0.003 [0.002]	0.021*** [0.007]	0.001 [0.002]	0.003 [0.002]
Inequality		1.473** [0.585]	1.764** [0.680]		1.245** [0.607]	0.985* [0.523]	1.983 [1.395]	1.058* [0.565]	0.992* [0.541]
Inequality (demeaned)				1.750** [0.830]					
Squared Inequality (demeaned)				-8.653 [8.438]					
Mean Suffrage	2.864	2.858	2.858	2.858	2.858	2.888	3.500	2.544	2.867
R-squared	0.626	0.657	0.621	0.618	0.647	0.755	0.660	0.620	0.781
Number of observations	296	231	231	231	231	221	76	155	178

Panel B: First-stage results

					2SLS (5)	2SLS (6)			
Crop Index					0.206*** [0.051]				
Inequality Lagged Two Periods						0.936*** [0.027]			
First-stage R-squared					0.632	0.896			
Kleibergen-Paap rank test for underidentification					11.025	78.646			
Test p-value					0.001	0.000			

Sources: See Appendix 1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.1$.

Table 4: Suffrage in the 13 colonies: five-year panel, with suffrage Dinkin as dependent variable

							OLS (7) -	OLS (8) -	OLS (9) - only

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Dinkin Lagged One Period	0.613***	0.591***	0.686***	0.690***	0.552***	0.712***	0.310*	0.739***	0.599***
	[0.072]	[0.103]	[0.090]	[0.088]	[0.134]	[0.065]	[0.163]	[0.085]	[0.081]
Labour	**1.952***	**2.068***			**2.79**	**1.293***	**2.166***	**1.435**	**1.600***
	[0.506]	[0.645]			[1.742]	[0.400]	[0.795]	[0.705]	[0.431]
Income		0.003*	0.004*	0.003	0.003*	0.002	0.012**	0.001	0.001
		[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.005]	[0.002]	[0.002]
Inequality		0.789*	1.126**		0.555	0.483	0.597	0.612	0.508
		[0.420]	[0.514]		[0.494]	[0.392]	[0.889]	[0.438]	[0.389]
Inequality (demeaned)				1.268*					
				[0.655]					
Squared Inequality (demeaned)				-6.598					
				[6.449]					
Mean Suffrage	2.432	2.515	2.515	2.515	2.515	2.553	2.855	2.348	2.636
R-squared	0.657	0.656	0.613	0.613	0.645	0.735	0.716	0.620	0.733
Number of observations	296	231	231	231	231	221	76	155	178

Panel B: First-stage results

Crop Index					0.195***				
					[0.051]				
Inequality Lagged Two Periods						0.932***			
						[0.028]			
First-stage R-squared					0.638	0.896			
Kleibergen-Paap rank test for underidentification					10.273	82.942			
Test p-value					0.001	0.000			

Sources: See Appendix 1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage Dinkin, which is similar to the variable Suffrage in Chart 1 and Table 3, but excludes restrictions for freemanship, formal patenting of lands, householding and being a "good person," and codes periods of no elections as the maximum number of restrictions in this restricted sample (6.0). See notes to Chart 1 and Appendix 1 on how the variable Suffrage is constructed. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Appendix 1: Data sources and description		
VARIABLE	**DESCRIPTION**	**SOURCE**
Dependent variables: Suffrage and Suffrage Dinkin	Constructed from qualitative descriptions for each colony's suffrage restrictions from McKinley (1905) coded as follows: • Restrictions that cannot be substituted with another restriction are coded as 1 • Restrictions that can be substituted with one other restriction are coded each as ¼ (so that the sum of the two restrictions is ½) • Restrictions that can be substituted with two other restrictions are coded each as 1/9 (so that the sum of the three restrictions is 1/3). • This coding is adopted in order to emphasise instances of suffrage liberalisation. Coding restrictions that can be substituted with another restriction as ½ and restrictions that can be substituted with two other restrictions as 1/3 produced very similar results. • Two types of restrictions indices are considered: • All the McKinley restrictions are summed up. These are the voting restrictions for: freemanship; particular status (e.g. the son of a freeholder); oath taking; freeholding; the existence of minimum freeholding; the existence of property or a minimum income; tax; residency; religion; race, having a family, having one's lands formally patented; householding; restrictions against felons, and being a "good person". The minimum number of restrictions in this sample is 1, the maximum is 6.5, and average values for different samples are presented in the regression tables. Periods of no elections are coded as having the maximum number of restrictions in the sample (6.5 restrictions). • For the robustness checks only: only the restrictions identified by Dinkin (1977) are considered (i.e., excluding restrictions for freemanship, formal patenting of lands, householding and being a "good person"). The minimum number of restrictions in this sample is 0, the maximum is 6, and periods of no elections are coded as having the maximum number of restrictions in the sample.	McKinley (1905) Dinkin (1977)
Labour	This variable was constructed by calculating the percentage of each colony's black population using data for the black and white population, for every 10 years. Exponential interpolation was used to fill in the missing values between the 10-year periods.	Historical Census Browser. Retrieved Aug 1, 2008, from the University of Virginia, Geospatial and Statistical Data Center: http://fisher.lib.virginia.edu/collecti ons/stats/histcensus/index.html
Income	This variable was constructed by coding narrative building trades wage data, which is measured in US cents per day. The conversion rate is assumed to be that 1 British shilling=73 cents. In colonial times, 1 British pound=20 shillings=12 pence. This is the series with the greatest number of observations in the original source. I also compared this series with data for the following occupations: agricultural labour, clothing trades, common labour, iron works and maritime occupations. In order to use the building trade wage series, I completed the following steps: • averaged the building trade wages by year and colony, ignoring occupations and excluding	Bureau of Labor Statistics (1929) "History of Wages in the United States from Colonial Times to 1928."

	observations where board was part of the wage • assumed that wages persisted for two years before and two years after a given year • filled up the missing wage observations by assigning the wage values for colonies from the same region (again assuming that this new wage observation persisted for two years before and two years after the given year) • The remaining missing values were filled by linear interpolation.	
Inequality	This variable is measured as the percentage of a colony's landless population. I assumed that inequality persisted for five years before and five years after a given year, as arguably inequality is more "sticky" than wages. I linked inequality to Suffrage as follows: • arranged the inequality data by colony, year and locality (county, city or state) • after a manual inspection, excluded observations where only one place per colony was available and the author did not explicitly mention that this place was representative for the colony as a whole • for a particular year and colony, looked for percentage landless data for (1) the colony and, if not available, (2) another colony in the region. • The remaining missing values were filled by linear interpolation.	Kulikoff (1986, 2000) Main (1965) Nash (1979)
Latitude	These data were obtained from Tiger Map through Census Bureau. I use the latitude for each colony's state capital.	Tiger Map (Census Bureau). http://www.census.gov/geo/www/tiger/latlng.txt (accessed 10 October, 2011)
Average Temperature	Average Temperature is measured in Fahrenheit, and is averaged over the period January 1895 - January 2008.	National Climatic Data Center, US Department of Commerce. http://www7.ncdc.noaa.gov/CDO/cdo (accessed 1 August, 2008)
Instrument for Labour: Crop Index	See text. This variable is measured in British pence per pound.	Prices: Purvis (1999), Historical Statistics of the US Weights: National Agricultural Statistics Service, United States Department of Agriculture. http://www.nass.usda.gov/Data_and_Statistics/Quick_Stats/index.asp (accessed 1 August, 2008). Qualitative weights (for Crop Index comparison): McCusker and Menard (1991)

Supplementary materials
Model

1 Assumptions

This is a two-period game, and there are two classes of agents based on income: rich (landowners) and poor (workers from Britain), where the income of the rich is above mean income, while the income of the poor is below mean income. There is no within-class heterogeneity. In the second period, there are also slaves who are perfect substitutes for poor workers. The number of poor is P, while the number of rich is R, with $R < P$. There is no capital in this economy. Output (Y) is produced according to a Cobb-Douglas production function with constant returns to scale: $Y = A(S+P)^{\alpha}L^{1-\alpha}$, where A is a fixed productivity parameter, S and P are the number of slaves and poor workers, respectively, L is land and $\alpha < 1$. The price of output is normalised to 1.

As explained in the text, the poor are paid their marginal product in a democracy and a subsistence wage in an autocracy. The most preferred tax rate of each agent is found by maximising her after-tax income, which is assumed to follow a standard Meltzer-Richard setup. More precisely, income is $(1-\tau)y_i + \tau y - \frac{\tau^2 y}{2}$, where τ is the tax rate, y_i is the pre-tax income of each group, y is mean income, and $i = r$ or p. After maximising this expression with respect to τ, I obtain that the poor prefer a tax rate of $\frac{y-y_p}{y}$, while the rich prefer a tax rate of 0. For simplicity, extending or not extending the suffrage is a binary decision (for example, only complete democracy or complete autocracy can be established). There is no discounting, and the utility of each agent is simply the sum of the payoffs from the two periods.[1] Chart S1 summarises the structure of the game.

2 Timing and payoffs

Period 1

Initially, this economy is populated only by rich agents, while all poor agents reside in Britain. Because of the specification of the production function, the income of the rich directly depends on the number of employed labourers. In order to attract poor British workers to the American colonies, the rich could simply set the wage equal to the marginal product of the number of poor migrants. However, because the poor know that the rich are likely to renege on their wage promise, they migrate only on the condition that inclusive representative institutions are provided as well. Suffrage regulations can be both economic (such as income, freeholding or residency) and non-economic (such as religion, race or family status) because elites can use both categories

---

[1]The insights of the model can also be extended to the infinite-horizon setting. The model can be re-worked so that slavery happens in some period $t + k$. In this case, the rich will weigh their life-time utilities of contracting the suffrage versus keeping a democratic regime, given that some fraction of the poor leaves forever after the contraction of the suffrage.

of restrictions to cater to groups on whose labour they depend. I also assume that voting yields additional non-monetary benefits to the poor.

The rich move first and decide whether to grant democracy or not. In an autocracy, only a limited number of poor come to the colonies, the game ends and the rich get a payoff of $\underline{y_r}$, or the lower bound of their income. If the rich grant the suffrage, a large number of poor come, and the payoff of the rich is:

$$V_r^1 = (1 - \tau_p)y_r + \tau_p y - \frac{\tau_p^2 y}{2} \tag{1}$$

where y_r is the income of each rich person, y is average income, and τ_p is the most preferred tax of the poor (for clarity, I omit the * when referring to the most preferred tax of the poor). I assume that in the first period the available supply of workers from Britain is sufficient to maximise the income of the rich if the suffrage is given to the poor. Since output is produced according to a Cobb-Douglas production function and the wage of the poor is equal to their marginal product, this expression is equivalent to:

$$V_r^1 = (1 - \tau_p)(1 - \alpha)\frac{Y(P)}{R} + \tau_p \frac{Y(P)}{R + P} - \frac{\tau_p^2}{2}\frac{Y(P)}{R + P}, \tag{2}$$

where all the variables are as defined above, and $Y(P) = A(P)^\alpha L^{1-\alpha}$.[2]

The poor, who can only migrate to the colonies in the first period, observe the behaviour of the rich and decide whether to move or not. If the poor stay in Britain, their after-tax income is y_p^{uk}. If the poor come to America without being given the right to vote, they get a payoff of $w - m_i$, where w is a subsistence wage that the poor earn in an autocracy (since the rich cannot commit to promises of a higher wage) and m_i is the cost of migrating ($i = h, l$, meaning that the cost of migration is either high or low, and it is disclosed to the poor before they migrate). If the poor come to the colonies and the rich decide to extend the suffrage, the payoff of the poor is:

$$V_p^1 = (1 - \tau_p)y_p + \tau_p y - \frac{\tau_p^2 y}{2} + \pi_p^{us} - m_i, \tag{3}$$

where the terms are as defined above, and π_p^{us} captures any non-monetary benefits of representative institutions in America.

Since the income of the poor consists of wages, which are equal to their marginal product, this expression is equivalent to:

$$V_p^1 = (1 - \tau_p)\alpha\frac{Y(P)}{P} + \tau_p \frac{Y(P)}{R + P} - \frac{\tau_p^2}{2}\frac{Y(P)}{R + P} + \pi_p^{us} - m_i. \tag{4}$$

Similarly, the most preferred tax of the poor is found to be:

---

[2]In a competitive labour market, the poor are paid their marginal product, while the rich retain the rest of the output. After differentiating the output equation with respect to P, $\frac{\partial Y}{\partial P} = \alpha A P^{\alpha-1} L^{1-\alpha} = \frac{\alpha Y}{P}$. Therefore, the rich get $(1 - \alpha)Y$ out of total output, while each rich person obtains $\frac{(1-\alpha)Y}{R}$.

$$\tau_p = \frac{y - y_p}{y} = 1 - \alpha\frac{R + P}{P} \tag{5}$$

To narrow down the number of cases to be analysed, I assume that it is optimal only for a limited number of poor to leave Britain without being granted representative institutions in the colonies. Similarly, I assume that it is always optimal for any poor person to leave Britain and stay in America for at least one period if they are given the right to vote. These two conditions are captured by the following inequalities:

Period 2

In the second period, two changes in labour markets take place. First, the rich can acquire slaves (denoted by S), who are perfect substitutes for poor migrants.[3] This "slavery shock" is exogenous, but the rich are uncertain about the value of the *total labour supply*, and this uncertainty determines whether they impose an autocracy or keep the democratic regime from the previous period. Second, if the rich decide to contract the suffrage, it is optimal for some of the poor to re-enter Britain (after paying a cost of migration, which for simplicity I assume to be the same as that paid when coming to America).[4] However, even with a contraction of the suffrage, a fraction θ of the poor stay as labourers, where θ is determined by the exogenous proportion of poor that have high costs of re-entry. The rich are uncertain about the exact "mixture" of available labour: either the supply of labour maximises their payoff, or it is such that their payoff is lower than that in the democratic case. This uncertainty can arise because of either θ or S. In particular, the rich believe that the labour supply is high ($\theta_h P + S_h$, providing them with a payoff higher than that under a democracy) with probability λ and low ($\theta_l P + S_l$, providing them with a payoff lower than that under a democracy) with probability $1 - \lambda$.

If the rich decide to change the political system, they implement their most preferred tax rate of 0, but also pay a fixed cost c, and lose a fraction $(1 - \theta)$ of the poor labourers. The term c captures any direct costs (such as repression or transaction costs) that the rich incur when changing the regime from a democratic to an authoritarian one. If the rich keep the democratic system, their payoff is the same as that under democracy in the first period. Since in a democracy the poor have a majority in the Assembly, the spread of slavery is blocked and the rich are forced to use the poor as labourers. As a result, in a democratic system the wages of the poor remain the same as those in the previous period (without slavery).

Conversely, instead of setting wages equal to marginal product, in an autocracy the rich can simply pay the poor some subsistence value w and thus appropriate a higher proportion of the output. The value of this "wage" is equivalent to the price and maintenance of a single slave. The

---

[3]For simplicity, I assume that slaves and poor labourers yield the same return for the rich if the tax rate is 0. Overall, slaves are cheaper because they do not demand any voting rights, and thus the rich can implement their most preferred tax rate.

[4]The current model can easily be adapted to take into account inter-colonial migration as well. For example, the outside option could be a weighted average of the income/representative institutions in Britain and those in the closest colony/or a colonial average.

expected payoff to the rich in an authoritarian regime is thus:

$$V_r^2 = \lambda V_r^{h2} + (1-\lambda)V_r^{l2} - c = \lambda\left[\frac{Y(\theta_h P + S_h)}{R} - w\frac{\theta_h P + S_h}{R}\right] + (1-\lambda)\left[\frac{Y(\theta_l P + S_l)}{R} - w\frac{\theta_l P + S_l}{R}\right] - c, \quad (7)$$

where $Y(\theta_h P + S_h) = A(\theta_h P + S_h)^\alpha L^{1-\alpha}$ and $Y(\theta_l P + S_l) = A(\theta_l P + S_l)^\alpha L^{1-\alpha}$.

After the rich decide whether to contract the suffrage or keep it the same, the poor choose to stay or not. If the rich decide not to change the suffrage, the payoff of each poor worker is the same as that in the first period, excluding the cost of migration m_i. If the rich decide to contract the suffrage, $(1-\theta)$ of the poor leave, getting an income of $y_p^{uk} - m_l$, while a fraction θ of the poor stay, and the payoff of the latter group is $V_p^{c2} = w$.

3 Equilibria

I show that this game has two subgame perfect Nash equilibria (SPNEs) depending on the value of λ.[5] In both equilbria, in the first period the rich will extend the suffrage and all the poor will come to the colonies. Starting with the second period, the first equilibrium involves the rich keeping the democratic system and all the poor staying. In the second equilibrium, the rich will contract the suffrage, and a fraction $(1-\theta)$ of the poor will leave. Below, I provide the conditions under which these equilibria exist.

Should the rich choose an autocratic regime in the second period, it is optimal only for those poor with low costs of migration to leave the colonies as long as $w \leq y_p^{uk} - m_l$, while it is best for the poor with high costs of migration to stay if $w \geq y_p^{uk} - m_h$. Therefore, the decision of the rich whether to contract the suffrage or to keep it the same will determine if *all* poor or only *some* poor stay in the colonies in the second period. In turn, in the second period the rich will weigh their utility of contracting the suffrage versus keeping it the same. Given that all the poor come in the first period, the rich only need to consider the impact of their action in the second period, which yields that the rich will contract the suffrage if:

$$\lambda\left[\frac{Y(\theta_h P + S_h)}{R} - w\frac{\theta_h P + S_h}{R}\right] + (1-\lambda)\left[\frac{Y(\theta_l P + S_l)}{R} - w\frac{\theta_l P + S_l}{R}\right] - c \geq$$
$$(1-\tau_p)(1-\alpha)\frac{Y(P)}{R} + \tau_p\frac{Y(P)}{R+P} - \frac{\tau_p^2}{2}\frac{Y(P)}{R+P} \quad (8)$$

From this expression, in this game the rich follow a *cutoff strategy*: they will contract the suffrage if $\lambda \geq \lambda^*$ and keep the suffrage the same if $\lambda \leq \lambda^*$, where:

$$\lambda^* = \frac{\left[(1-\tau_p)(1-\alpha)\frac{Y(P)}{R} + \tau_p\frac{Y(P)}{R+P} - \frac{\tau_p^2}{2}\frac{Y(P)}{R+P}\right] - \left[\frac{Y(\theta_l P + S_l)}{R} - w\frac{\theta_l P + S_l}{R} - c\right]}{\left[\frac{Y(\theta_h P + S_h)}{R} - w\frac{\theta_h P + S_h}{R}\right] - \left[\frac{Y(\theta_l P + S_l)}{R} - w\frac{\theta_l P + S_l}{R}\right]}. \quad (9)$$

---

[5]Because of the assumption that very few poor will come in an autocracy, there is a trivial Nash equilibrium in which the rich do not extend the suffrage and only a limited number of poor come in the first period. However, this is not an SPNE because the rich are better off giving the poor the right to vote.

Equation (9) reveals that when choosing representative institutions in the second period, the rich take into account both inequality and the state of the labour market. Since average income in this economy is $\frac{Y(P)}{R+P} = \alpha\frac{Y(P)}{P}\frac{P}{R+P} + (1-\alpha)\frac{Y(P)}{R}\frac{R}{R+P}$, fixing R and P, α represents the income share of the poor and is thus a measure of inequality (as α increases, inequality decreases). In order to study how λ^* varies with inequality, I differentiate (9) with respect to α, after first substituting expression (5) for τ_p, yielding that:[6]

$$\frac{\partial\lambda^*}{\partial\alpha} = \frac{Y(P)}{P^2R}[(1-2\alpha)(R+P)P - \alpha R(R+P)]. \tag{10}$$

This expression is positive for $\alpha < \frac{P}{2P+R}$, and it is negative for $\alpha > \frac{P}{2P+R}$. In addition, since by assumption $y_r < y < y_p$ and output is produced according to a Cobb-Douglas production function, we need that $\alpha\frac{Y(P)}{P} < \frac{Y(P)}{R+P} < (1-\alpha)\frac{Y(P)}{R}$ or that $\alpha < \frac{P}{R+P}$. Therefore, for small values of α (when $\alpha < \frac{P}{2P+R}$) the derivative is in line with the linear redistribution model: an increase in inequality makes it less likely for democracy to emerge. When $\frac{P}{R+P} > \alpha > \frac{P}{2P+R}$, the opposite effect kicks in - an increase in inequality makes a democratic government more likely. In the empirical section in the text, I find no evidence for a non-linear effect of inequality on suffrage.

More importantly, expression (9) also captures the importance of labour markets.[7] From above, $V_r^{h2} = \frac{Y(\theta_h P + S_h)}{R} - w\frac{\theta_h P + S_h}{R}$ and $V_r^{l2} = \frac{Y(\theta_l P + S_l)}{R} - w\frac{\theta_l P + S_l}{R}$ (note that this setup ensures that the values of V_r^{h2} and V_r^{l2} may change because of changes in either θ or S, but not because of α). As the difference $V_r^{h2} - V_r^{l2}$ increases, meaning that the income of the rich under a labour surplus becomes higher *relative* to that under a labour shortage, λ^* decreases and the rich are more likely to contract the suffrage (and vice versa). In particular, if an increase in S_h offsets a decrease in θ_h (such as in a period of rapid influx of slaves accompanied by a lower supply of indentured servants), then a tightening of the suffrage will occur.

4 The role of the middle class

The historical evidence suggests that the introduction of slavery led to an increase in the number of small and medium farmers in the South not only because of the the cultivation of new frontier land, but also because of the burgeoning Southern land market (see Williams 1969). While initially land transactions only involved proprietors, their friends and speculators, after 1700 Southern elites sold a large amount of land to the general public. For example, during the first half of the 18th century, approximately a quarter to two-fifths of the new settlers in Southside Virginia acquired their land (ranging in size from 270 to 300 acres) from other owners (Kulikoff 2000, p. 154). In South Carolina, the frequency of land sales went up in the period 1720-75, and the size of the median land sale was 300 acres (Ryden and Menard 2005). Similarly, in the period

---

[6]Note that I assume that overall income and average income do not change as α changes. This simplifies the analysis and allows me not to consider the effect of the terms in the denominator of expression (9) when calculating the derivative in equation (10).

[7]I assume that $\frac{Y(\theta_l P + S_l)}{R} - w\frac{\theta_l P + S_l}{R} - c \to 0$. In other words, the rich would never have an incentive to change the political system if they knew for sure that the labour market was tight.

1738-78 the Lords Baltimore sold a large amount of isolated Maryland land.[8]

I assume that the middle class consists of small landowners that work their own farms, meaning that the total supply of slaves and white workers does not affect their income.[9] While the existence of a middle class should not change the incentives of the rich to choose the democratic equilibrium, it may be more profitable for them to impose a partial democracy instead of an autocracy even if the expected value of the labour supply is high. First, elites may prefer to give voting rights to the middle class, as opposed to implementing a full autocracy, if the rich need to create an alliance in order to counteract any resistance threats from the poor and/or the slaves. In addition, the tax rate in a democratic regime that includes only the middle class and the rich is lower than the one implemented when the poor are given the suffrage. Such a tax rate should be particularly attractive to elites in cases of labour surplus and/or high inequality, and also in cases of low middle class-rich inequality. Finally, once the suffrage is restricted to include only the middle class, the Assembly can pass legislation that benefits only these two groups. The idea is that without the interference of the poor, the interests of the rich and the middle class are more closely aligned, and this produces representative institutions that increase the income of both groups, without additional taxation.[10] In short, in cases of sufficient labour supply, the existence of a middle class may provide incentives to the rich to implement a partial democracy instead of an autocratic regime.

---

[8]The difficulty of overseeing slaves on large tracts of land and expectations of economic growth in the South could explain the surge in Southern land sales after 1700.

[9]Consistent with the historical narrative, I assume that the relative sizes of the three classes are such that the median voter in a perfect democracy is still a poor individual.

[10]An example of a policy preferred only by the poor could be public education. While the rich and the middle class may prefer private education and thus may not be willing to spend tax money on providing universal education, these two groups will not be able to implement their preferred policy if the poor are given the right to vote.

Table S1A: Characteristics of the 13 British American colonies

Colony	Narrow region	Broad region	Date first settled (by the British)	Average latitude	Type of initial settlement	Main agricultural products
Massachusetts	North	North	1629	42.336	Charter	fish, livestock, grains
Connecticut	North	North	1662	41.766	Charter	fish, livestock, grains
Rhode Island	North	North	1663	41.822	Charter	fish, livestock, grains
New Hampshire	North	North	1680	43.232	Royal	fish, livestock, grains
New York	Middle	North	1664	42.666	Proprietary	wheat
New Jersey	Middle	North	1664	40.223	Proprietary	wheat
Delaware	Middle	North	1681	39.159	Proprietary	wheat
Pennsylvania	Middle	North	1681	40.276	Proprietary	wheat
Virginia	Upper South	South	1606	37.531	Charter	tobacco, wheat
Maryland	Upper South	South	1632	38.972	Proprietary	tobacco, wheat
North Carolina	Lower South	South	1663	35.822	Proprietary	tobacco, wheat
South Carolina	Lower South	South	1663	34.039	Proprietary	rice, indigo
Georgia	Lower South	South	1732	33.762	Proprietary	rice, indigo

Sources: Purvis (1999), McCusker and Menard (1991).

Notes: This table summarises geographical and settlement characteristics of the 13 British American colonies. Maine is part of the main data set, because McKinley (1905) provides Suffrage data for Maine for the period 1639-51. However, Maine is not included in this table because it was an enclave of Massachusetts until 1820.

Table S1B: Colonial British America: characteristics of main agricultural exports

Crop	Origin	Conditions required for growing	Economies of scale	Year production started	Main region exporting crop	Share of crop in value of region's exports (1768-1772)
Rice	East India to South China	Tropical and sub-tropical climate around the shores, requires irrigation, prefers clay loam soil.	strong	1695	Lower South	55.40%
Tobacco	South America	Highest valued tobacco is grown in tropical and semi-tropical climates. The type of soil determines the texture, whereas climate determines the aroma.	medium	1612	Upper South	72.20%
Wheat	Near East; Georgia	Can be grown in almost any type of climate and over a great variety of soils.	none	from settlement	Middle	72.10%

Sources: Bailey (1907), Engerman, Sutch and Wright (2006), Fogel (1989), Gray (1933), Harlan (1992), and McCusker and Menard (1991).

Notes: This table lists the three most important agricultural products exported by the 13 British American colonies. The largest export of New England as of 1768-1772 was fish, but New Englanders also grew large amounts of wheat until the Middle colonies specialised in wheat growing in the late 1600s. New England was also suitable for the growing of cigar wrappers, but opposition from Britain prevented the development of this industry in colonial times.

Table S2: Suffrage in the 13 colonies: regional panel, with the ratio of Caribbean slave prices over the crop index as measure of Labour

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)	OLS (7)	OLS (8)	OLS (9)	OLS (10)	OLS (11)	OLS (12)
Suffrage Lagged One Period	0.696*** [0.061]	0.657*** [0.059]	0.676*** [0.052]	0.825*** [0.050]	0.808*** [0.050]	0.803*** [0.045]	0.685*** [0.156]	0.634*** [0.157]	0.681*** [0.131]	0.847*** [0.082]	0.829*** [0.086]	0.850*** [0.071]
Caribbean Slave Prices/Crop Index	-0.004** [0.002]	-0.003* [0.002]		-0.002 [0.001]	-0.002 [0.001]		-0.004 [0.004]	-0.003 [0.004]		-0.002 [0.002]	-0.002 [0.002]	
Income		0.006** [0.003]	0.010*** [0.003]		0.003 [0.002]	0.005** [0.002]		0.008 [0.007]	0.011 [0.007]		0.002 [0.004]	0.003 [0.004]
Inequality		2.029*** [0.540]	2.921*** [0.533]		0.994** [0.426]	1.594*** [0.479]		2.037 [1.235]	1.995 [1.226]		1.03 [0.775]	1.006 [0.768]
Mean Suffrage	3.100	3.100	3.053	3.093	3.093	3.050	3.119	3.119	3.119	3.071	3.071	3.071
R-squared	0.78	0.797	0.738	0.862	0.866	0.806	0.775	0.794	0.791	0.862	0.866	0.865
Number of observations	192	192	235	196	196	239	40	40	40	66	66	66

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained in the following way. Columns 1-3: Annual panel, five-year lags; Columns 4-6: Annual panel, three-year lags; Columns 7-9: Panel obtained by taking every fifth year for each colony; Columns 10-12: Panel obtained by taking every third year for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. Caribbean Slave Prices/Crop Index is the ratio of Caribbean Slave Prices (measured in pounds sterling) and the Crop Index (measured in English pence per pound). All independent variables are lagged by one period, which is three or five years depending on the panel. All regressions control for Year and Latitude. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S3: Suffrage in the 13 colonies: five-year panel, with binary indicator as measure of Labour

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only after 1700
Suffrage Lagged One Period	0.627***	0.612***	0.671***	0.683***	0.588***	0.749***	0.322*	0.706***
	[0.068]	[0.091]	[0.078]	[0.075]	[0.111]	[0.063]	[0.176]	[0.083]
Binary Indicator	**0.569***	**0.633**			**0.849**	**0.249**	**0.089**	**0.281***
	[0.204]	[0.249]			[0.600]	[0.153]	[0.400]	[0.159]
Income		0.005**	0.005*	0.004	0.005**	0.003	0.023***	0.003
		[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.008]	[0.002]
Inequality		1.178**	1.764**		0.911	0.921*	2.052	0.448
		[0.546]	[0.680]		[0.738]	[0.494]	[1.611]	[0.544]
Inequality (demeaned)				1.750**				
				[0.830]				
Squared Inequality (demeaned)				-8.653				
				[8.438]				
Mean Suffrage	2.852	2.858	2.858	2.858	2.858	2.888	3.500	2.867
R-squared	0.585	0.639	0.621	0.618	0.636	0.745	0.611	0.765
Number of observations	298	231	231	231	231	221	76	178

Panel B: First-stage results

Crop Index					0.771***			
					[0.093]			
Inequality Lagged Two Periods						0.934***		
						[0.031]		
First-stage R-squared					0.696	0.896		
Kleibergen-Paap rank test for underidentification					37.199	83.506		
Test p-value					0.000	0.000		

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth year for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. Binary Indicator is a variable that is 1 for the South after 1700 and 0 for the South before 1700 and for the North. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S4: Suffrage in the 13 colonies: 10-year panel

					Panel A: Second-stage results				
Independent variables	**OLS (1)**	**OLS (2)**	**OLS (3)**	**OLS (4)**	**2SLS (5)**	**2SLS (6)**	**OLS (7) - only South**	**OLS (8) - only North**	**OLS (9) - only after 1700**
Suffrage Lagged One Period	0.445***	0.363***	0.472***	0.486***	0.303*	0.441***	0.091	0.533***	0.359**
	[0.084]	[0.126]	[0.124]	[0.122]	[0.171]	[0.135]	[0.138]	[0.161]	[0.145]
Labour	**3.445*****	**3.400*****			**4.962**	**2.596*****	**3.717****	**3.331****	**3.059*****
	[0.860]	[0.968]			[3.712]	[0.950]	[1.522]	[1.635]	[1.021]
Income		0.004	0.006	0.005	0.003	0.003	0.017**	0.001	0.003
		[0.003]	[0.004]	[0.004]	[0.004]	[0.004]	[0.007]	[0.003]	[0.003]
Inequality		2.158**	2.820**		1.636	2.463**	2.767	1.740*	2.186**
		[0.947]	[1.127]		[1.215]	[1.046]	[2.660]	[1.012]	[1.038]
Inequality (demeaned)				2.955**					
				[1.312]					
Squared Inequality (demeaned)				-12.839					
				[12.558]					
Mean Suffrage	2.902	2.924	2.924	2.924	2.924	2.947	3.625	2.574	2.866
R-squared	0.463	0.596	0.514	0.508	0.568	0.658	0.598	0.519	0.685
Number of observations	144	108	108	108	108	98	36	72	88
					Panel B: First-stage results				
Crop Index					0.165*				
					[0.084]				
Inequality Lagged Two Periods						0.816***			
						[0.056]			
First-stage R-squared					0.568	0.658			
Kleibergen-Paap rank test for underidentification					3.954	29.956			
Test p-value					0.047	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every 10th observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. All independent variables are lagged by one period (10 years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S5: Suffrage in the 13 colonies: five-year panel, averages

							Panel A: Second-stage results		

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.734***	0.756***	0.819***	0.824***	0.728***	0.745***	0.663***	0.768***	0.745***
	[0.050]	[0.056]	[0.045]	[0.044]	[0.073]	[0.062]	[0.093]	[0.069]	[0.083]
Labour	**1.524***	**1.460***			**2.053**	**1.397***	**1.523**	**1.408*	**1.520***
	[0.427]	[0.408]			[1.251]	[0.421]	[0.664]	[0.787]	[0.496]
Income		0.001	0.001	0.001	0.001	0.001	0.008*	-0.001	0.002
		[0.001]	[0.002]	[0.002]	[0.001]	[0.001]	[0.004]	[0.001]	[0.001]
Inequality		0.875*	1.003**		0.77	1.166**	0.672	0.857	0.595
		[0.468]	[0.498]		[0.475]	[0.545]	[0.769]	[0.560]	[0.443]
Inequality (demeaned)				0.959*					
				[0.531]					
Squared Inequality (demeaned)				-4.787					
				[4.807]					
Mean Suffrage	2.879	2.888	2.888	2.888	2.888	2.884	3.524	2.580	2.868
R-squared	0.703	0.793	0.779	0.778	0.790	0.793	0.835	0.706	0.865
Number of observations	298	239	239	239	239	229	78	161	178

Panel B: First-stage results

Crop Index					0.207***				
					[0.049]				
Inequality Lagged Two Periods						0.939***			
						[0.022]			
First-stage R-squared					0.639	0.917			
Kleibergen-Paap rank test for underidentification					13.586	86.488			
Test p-value					0.000	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking five-year averages of all variables. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S6: Suffrage in the 13 colonies: five-year panel, with suffrage obtained through principal component analysis

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.561*** [0.079]	0.482*** [0.121]	0.658*** [0.095]	0.679*** [0.090]	0.044 [0.321]	0.618*** [0.111]	0.408** [0.162]	0.587*** [0.106]	0.642*** [0.123]
Labour	**2.285*** [0.612]**	**2.879*** [0.895]**			**9.603** [3.869]**	**2.009** [0.815]**	**3.103** [1.476]**	**2.524*** [0.840]**	**1.778* [0.935]**
Income		0.005* [0.002]	0.005** [0.002]	0.005* [0.002]	0.003 [0.003]	0.004* [0.002]	0.014* [0.008]	0.002 [0.001]	0.004* [0.002]
Inequality		0.904** [0.404]	1.178** [0.511]		0.183 [0.480]	0.426 [0.363]	2.22 [1.454]	0.485 [0.323]	0.763 [0.480]
Inequality (demeaned)				1.108* [0.643]					
Squared Inequality (demeaned)				-9.771 [6.680]					
Mean Suffrage	-0.495	-0.583	-0.583	-0.583	-0.583	-0.576	0.416	-1.073	-0.505
R-squared	0.676	0.771	0.724	0.723	0.514	0.810	0.679	0.563	0.838
Number of observations	296	231	231	231	231	221	76	155	178

Panel B: First-stage results

Crop Index					0.139** [0.060]				
Inequality Lagged Two Periods						0.921*** [0.029]			
First-stage R-squared					0.705	0.895			
Kleibergen-Paap rank test for underidentification					3.958	82.257			
Test p-value					0.047	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is obtained by running principal component analysis on the Suffrage restrictions which are part of the variable Suffrage (from Table 3 and Chart 1), where the factor with the biggest eigenvalue is used in the analysis. See notes to Chart 1 and Appendix 1 on how the variable Sufrage is constructed. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S7: Suffrage in the 13 colonies: annual panel, using Prais-Winsten specification

Independent variables	GLS (1)	GLS (2)	GLS (3)	GLS (4)	GLS (5) - only South	GLS (6) - only North	GLS (7) - only after 1700
Suffrage Lagged One Period	0.921***	0.908***	0.929***	0.931***	0.884***	0.908***	0.921***
	[0.020]	[0.027]	[0.022]	[0.021]	[0.044]	[0.036]	[0.027]
Labour	**0.496***	**0.500***			**0.480*	**0.433*	**0.457**
	[0.160]	[0.176]			[0.256]	[0.239]	[0.179]
Income		0.001	0.001	0	0.003*	0	0.001
		[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	[0.001]
Inequality		0.205	0.201		-0.127	0.241	0.135
		[0.204]	[0.204]		[0.282]	[0.288]	[0.129]
Inequality (demeaned)				0.1			
				[0.218]			
Squared Inequality (demeaned)				-1.008			
				[1.799]			
Mean Suffrage	2.927	2.924	2.924	2.924	3.549	2.582	2.881
R-squared	0.905	0.904	0.903	0.903	0.914	0.869	0.943
Number of observations	1155	1041	1041	1041	368	673	816

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The regressions use a Prais-Winsten specification, which employs the GLS method and assumes that the errors have an AR (1) structure. The data set is an annual panel that drops the observations described in Section 4. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. All independent variables are lagged by one period (one year). All regressions control for Year and Latitude. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S8: Suffrage in the 13 colonies: five-year panel, using Tobit specification

Independent variables	Tobit (1)	Tobit (2)	Tobit (3)	Tobit (4)	Tobit (5) - only South	Tobit (6) - only North	Tobit (7) - only after 1700
Suffrage Lagged One Period	0.695***	0.596***	0.682***	0.694***	0.252	0.699***	0.648***
	[0.068]	[0.095]	[0.079]	[0.076]	[0.180]	[0.097]	[0.091]
Labour	**1.938***	**2.312***			**2.349**	**1.501*	**1.710***
	[0.644]	[0.746]			[1.022]	[0.866]	[0.613]
Income		0.004*	0.005*	0.004	0.022***	0.001	0.003
		[0.002]	[0.002]	[0.002]	[0.007]	[0.002]	[0.002]
Inequality		1.481**	1.770**		2.029	1.076*	0.982*
		[0.596]	[0.693]		[1.478]	[0.561]	[0.537]
Inequality (demeaned)				1.784**			
				[0.856]			
Squared Inequality (demeaned)				-8.952			
				[8.825]			
Mean Suffrage	3.044	2.942	2.942	2.942	3.754	2.574	2.888
Nagel-Kerke R-squared	0.627	0.660	0.623	0.620	0.640	0.652	0.808
Number of observations	296	231	231	231	76	155	178

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from Tobit regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. All independent variables are lagged by one period (five years). All regressions control for Year and Latitude. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S9: Suffrage in the 13 colonies: five-year panel, including exports

Independent variables	Main Spec. (1)	Suffrage Dinkin (2)	Using PCSEs (3)	Ethnic Frac. (4)	Religious Frac. (5)	Urbanization & Density (6)	Proprietary & Charter (7)	Tobit (8)
Suffrage Lagged One Period	0.649***	0.636***	0.649***	0.645***	0.638***	0.641***	0.631***	0.658***
	[0.094]	[0.088]	[0.067]	[0.095]	[0.099]	[0.100]	[0.097]	[0.094]
Labour	**1.541****	**1.365****	**1.541****	**1.426****	**1.496****	**1.442****	**2.124****	**1.519****
	[0.667]	[0.488]	[0.597]	[0.622]	[0.657]	[0.681]	[0.923]	[0.672]
Income	0.003	0.001	0.003	0.003	0.002	0.002	0.002	0.003
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Inequality	0.203	-0.157	0.203	0.125	0.313	0.376	0.093	0.071
	[0.554]	[0.486]	[0.732]	[0.548]	[0.561]	[0.589]	[0.562]	[0.599]
Exports	0.054	0.005	0.054	0.033	0.04	0.069	0.057	0.063
	[0.110]	[0.090]	[0.099]	[0.108]	[0.109]	[0.121]	[0.103]	[0.113]
Ethnic Fractionalisation				-0.3				
				[0.246]				
Religious Fractionalisation					-0.381			
					[0.232]			
Urbanisation						0.523		
						[0.663]		
Population Density						0.004		
						[0.007]		
Proprietary Colony							0.300*	
							[0.177]	
Charter Colony							0.064	
							[0.152]	
Mean Suffrage	3.020	2.702	3.020	3.020	3.020	3.020	3.020	3.085
R-squared/Nagel-Kerke R-squared	0.743	0.712	0.743	0.745	0.746	0.746	0.751	0.759
Number of observations	160	160	160	160	160	160	160	160

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. See Tables S10, S11, S12 and S13 and the text for the definitions of Urbanisation, Population Density, the dummies for Proprietary and Charter Colony, Ethnic Fractionalisation and Religious Fractionalisation. Column 3 presents a regression with panel-corrected standard errors, while column 8 presents results from a Tobit specification. All independent variables are lagged by one period (five years). All regressions control for Year and Latitude. Robust standard errors are in parentheses (except in column 3). ***p<=0.01, **p<=0.05, *p<=0.1.

Table S10: Suffrage in the 13 colonies: five-year panel, with urbanisation and population density

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.645***	0.574***	0.646***	0.657***	0.435***	0.705***	0.223	0.665***	0.634***
	[0.064]	[0.095]	[0.083]	[0.080]	[0.139]	[0.073]	[0.168]	[0.106]	[0.098]
Labour	1.999***	2.178***			6.409**	1.335***	2.195	0.657	1.573***
	[0.585]	[0.702]			[3.010]	[0.513]	[1.513]	[0.860]	[0.585]
Income		0.004*	0.003	0.002	0.004	0.002	0.022***	0.001	0.002
		[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.007]	[0.002]	[0.002]
Inequality		1.383**	1.626**		0.91	0.912*	1.681	0.804	1.115*
		[0.569]	[0.638]		[0.614]	[0.509]	[1.577]	[0.504]	[0.585]
Inequality (demeaned)				1.586**					
				[0.783]					
Squared Inequality (demeaned)				-9.599					
				[8.711]					
Urbanisation	0.132	0.641**	0.947***	0.975***	0.046	0.42	0.304	0.662*	0.703
	[0.284]	[0.305]	[0.359]	[0.367]	[0.510]	[0.293]	[5.167]	[0.337]	[0.630]
Population Density	0.001	0.004	-0.003	-0.003	0.017	0.006	0.046	0.003	0.006
	[0.006]	[0.007]	[0.007]	[0.007]	[0.010]	[0.007]	[0.042]	[0.007]	[0.006]
Mean Suffrage	2.864	2.858	2.858	2.858	2.858	2.888	3.500	2.544	2.867
R-squared	0.627	0.664	0.634	0.632	0.553	0.759	0.666	0.631	0.785
Number of observations	296	231	231	231	231	221	76	155	178

Panel B: First-stage results

Crop Index					0.189***				
					[0.057]				
Inequality Lagged Two Periods						0.935***			
						[0.027]			
First-stage R-squared					0.693	0.896			
Kleibergen-Paap rank test for underidentification					5.848	78.465			
Test p-value					0.016	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. The table includes controls for urbanisation and population density. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S11: Suffrage in the 13 colonies: five-year panel, with dummies for proprietary and charter colony

Panel A: Second-stage results

Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.618*** [0.067]	0.537*** [0.096]	0.643*** [0.081]	0.657*** [0.078]	0.406*** [0.150]	0.679*** [0.073]	0.224 [0.177]	0.644*** [0.105]	0.619*** [0.098]
Labour	**2.581*** [0.684]**	**3.021*** [0.932]**			**6.730* [3.444]**	**1.955** [0.765]**	**2.906** [1.182]**	**2.275** [1.086]**	**2.417*** [0.906]**
Income		0.004** [0.002]	0.005** [0.002]	0.004* [0.002]	0.003 [0.002]	0.003 [0.002]	0.020*** [0.006]	0.002 [0.002]	0.003 [0.002]
Inequality		1.541*** [0.561]	1.740*** [0.640]		1.296** [0.567]	1.126** [0.505]	1.477 [1.528]	1.054* [0.547]	1.025* [0.549]
Inequality (demeaned)				1.682** [0.779]					
Squared Inequality (demeaned)				-6.666 [7.741]					
Proprietary Colony	0.300** [0.119]	0.258* [0.144]	-0.108 [0.103]	-0.116 [0.105]	0.708* [0.417]	0.216 [0.143]	0.401 [0.278]	0.174 [0.163]	0.303* [0.156]
Charter Colony	0.296** [0.145]	0.433*** [0.156]	0.261* [0.141]	0.241* [0.140]	0.645** [0.253]	0.254** [0.129]		0.336** [0.154]	0.098 [0.152]
Mean Suffrage	2.864	2.858	2.858	2.858	2.858	2.888	3.500	2.544	2.867
R-squared	0.638	0.675	0.631	0.627	0.608	0.762	0.677	0.637	0.790
Number of observations	296	231	231	231	231	221	76	155	178

Panel B: First-stage results

					2SLS (5)	2SLS (6)			
Crop Index					0.144** [0.059]				
Inequality Lagged Two Periods						0.942*** [0.027]			
First-stage R-squared					0.755	0.901			
Kleibergen-Paap rank test for underidentification					4.096	77.599			
Test p-value					0.043	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. The table includes dummies for proprietary and charter colonies, with the omitted category royal colony (see Table S1A and Section 2 in the text for more information). All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S12: Suffrage in the 13 colonies: five-year panel, with ethnic fractionalisation

Panel A: Second-stage results									
Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.641*** [0.064]	0.578*** [0.093]	0.670*** [0.078]	0.686*** [0.074]	0.531*** [0.117]	0.712*** [0.068]	0.229 [0.168]	0.679*** [0.099]	0.648*** [0.091]
Labour	**1.934*** [0.539]**	**2.177*** [0.654]**			**3.296* [1.802]**	**1.316*** [0.488]**	**2.254** [0.908]**	**2.510** [1.167]**	**1.631*** [0.578]**
Income		0.004** [0.002]	0.005* [0.002]	0.004 [0.002]	0.004** [0.002]	0.003 [0.002]	0.018*** [0.006]	0.002 [0.002]	0.003 [0.002]
Inequality		1.378** [0.568]	1.786*** [0.669]		1.168** [0.591]	0.919* [0.517]	1.567 [1.423]	1.097* [0.580]	0.944* [0.554]
Inequality (demeaned)				1.733** [0.834]					
Squared Inequality (demeaned)				-8.485 [8.774]					
Ethnic Fractionalisation	0.313 [0.271]	0.42 [0.275]	-0.013 [0.246]	-0.125 [0.244]	0.643 [0.424]	0.145 [0.199]	-3.232 [2.440]	0.376 [0.294]	-0.087 [0.217]
Mean Suffrage	2.864	2.858	2.858	2.858	2.858	2.888	3.500	2.544	2.867
R-squared	0.627	0.658	0.621	0.616	0.648	0.756	0.669	0.623	0.781
Number of observations	296	231	231	231	231	221	76	155	178
Panel B: First-stage results									
Crop Index					0.233*** [0.050]				
Inequality Lagged Two Periods						0.947*** [0.026]			
First-stage R-squared					0.641	0.9			
Kleibergen-Paap rank test for underidentification					11.576	82.494			
Test p-value					0.001	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. Ethnic Fractionalisation is a Herfindahl-Hirschman Index calculated from colony-level data on the ancestral origins of the white population in 1790. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

Table S13: suffrage in the 13 colonies: five-year panel, with religious fractionalisation

							Panel A: Second-stage results		
Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)	OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Suffrage Lagged One Period	0.640***	0.587***	0.671***	0.685***	0.485***	0.717***	0.228	0.692***	0.648***
	[0.065]	[0.093]	[0.078]	[0.074]	[0.158]	[0.070]	[0.169]	[0.099]	[0.093]
Labour	**2.108*****	**2.306*****			**5.069**	**1.361*****	**2.230****	**0.944**	**1.688*****
	[0.604]	[0.727]			[3.465]	[0.518]	[0.967]	[1.096]	[0.612]
Income		0.004*	0.004*	0.003	0.004*	0.003	0.018***	0.001	0.003
		[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.006]	[0.002]	[0.002]
Inequality		1.497**	1.806***		1.127	1.021*	1.554	1.029*	1.015*
		[0.595]	[0.692]		[0.688]	[0.535]	[1.420]	[0.554]	[0.549]
Inequality (demeaned)				1.788**					
				[0.853]					
Squared Inequality (demeaned)				-10.77					
				[9.974]					
Religious Fractionalisation	-0.041	-0.09	-0.167	-0.218	0.001	-0.088	-1.225	-0.165	-0.102
	[0.212]	[0.210]	[0.225]	[0.279]	[0.265]	[0.183]	[0.816]	[0.310]	[0.178]
Mean Suffrage	2.877	2.858	2.858	2.858	2.858	2.888	3.500	2.544	2.867
R-squared	0.622	0.657	0.621	0.618	0.606	0.755	0.670	0.620	0.781
Number of observations	294	231	231	231	231	221	76	155	178

							Panel B: First-stage results		
Crop Index					0.204***				
					[0.050]				
Inequality Lagged Two Periods						0.949***			
						[0.028]			
First-stage R-squared					0.656	0.899			
Kleibergen-Paap rank test for underidentification					4.029	80.439			
Test p-value					0.045	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set is obtained by taking every fifth observation for each colony. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. Religious Fractionalisation is a Herfindahl-Hirschman Index calculated from colony-level data on the distribution of different types of churches in 1750. There are no religious fractionalisation data for Maine. All independent variables are lagged by one period (five years). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.

20

Table S14: Suffrage in the 13 colonies: annual panel, including all colonies and observations

							OLS (7) - only South	OLS (8) - only North	OLS (9) - only after 1700
Panel A: Second-stage results									
Independent variables	OLS (1)	OLS (2)	OLS (3)	OLS (4)	2SLS (5)	2SLS (6)			
Suffrage Lagged One Period	0.906***	0.903***	0.913***	0.914***	0.906***	0.900***	0.887***	0.896***	0.922***
	[0.017]	[0.024]	[0.022]	[0.021]	[0.027]	[0.025]	[0.042]	[0.036]	[0.025]
Labour	**0.355***	**0.396***			**0.236**	**0.416***	**0.178**	**0.288**	**0.295**
	[0.138]	[0.151]			[0.412]	[0.157]	[0.157]	[0.225]	[0.139]
Income		0.001	0.001	0.001	0.001	0.001	0.002	0	0
		[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.002]	[0.000]	[0.001]
Inequality		0.267	0.306*		0.289*	0.294*	0.105	0.288	0.098
		[0.165]	[0.173]		[0.159]	[0.172]	[0.344]	[0.189]	[0.118]
Inequality (demeaned)				0.286					
				[0.185]					
Squared Inequality (demeaned)				-1.501					
				[1.694]					
Mean Suffrage	3.049	2.981	2.981	2.981	2.981	2.972	3.678	2.606	2.941
R-squared	0.868	0.883	0.881	0.881	0.882	0.880	0.877	0.847	0.934
Number of observations	1648	1267	1267	1267	1267	1256	443	824	954
Panel B: First-stage results									
Crop Index					0.219***				
					[0.024]				
Inequality Lagged Two Periods						0.987***			
						[0.005]			
First-stage R-squared					0.555	0.982			
Kleibergen-Paap rank test for underidentification					63.153	536.812			
Test p-value					0.000	0.000			

Sources: See Appendices 1 and S1 and the text for variable definitions and data sources.

Notes: This table shows results from OLS and 2SLS regressions explaining the evolution of the right to vote in the 13 British American colonies from 1606 to 1775. The data set includes observations for all colonies and years. The dependent variable is Suffrage. See notes to Chart 1 and Appendix 1 on how this variable is constructed. All independent variables are lagged by one period (1 year). Column 5 presents results from a 2SLS regression where Labour is instrumented with Crop Index, which is calculated by multiplying the propensity of each colony to produce tobacco, rice or wheat with the respective price of each crop. Column 6 presents results from a 2SLS regression where Inequality is instrumented with Inequality Lagged Two Periods. All regressions control for Year and all regressions, except the one in column 5, include controls for Latitude. In order to obtain stronger first-stage results, the regression in column 5 controls for Average Temperature. Robust standard errors are in parentheses. ***p<=0.01, **p<=0.05, *p<=0.1.
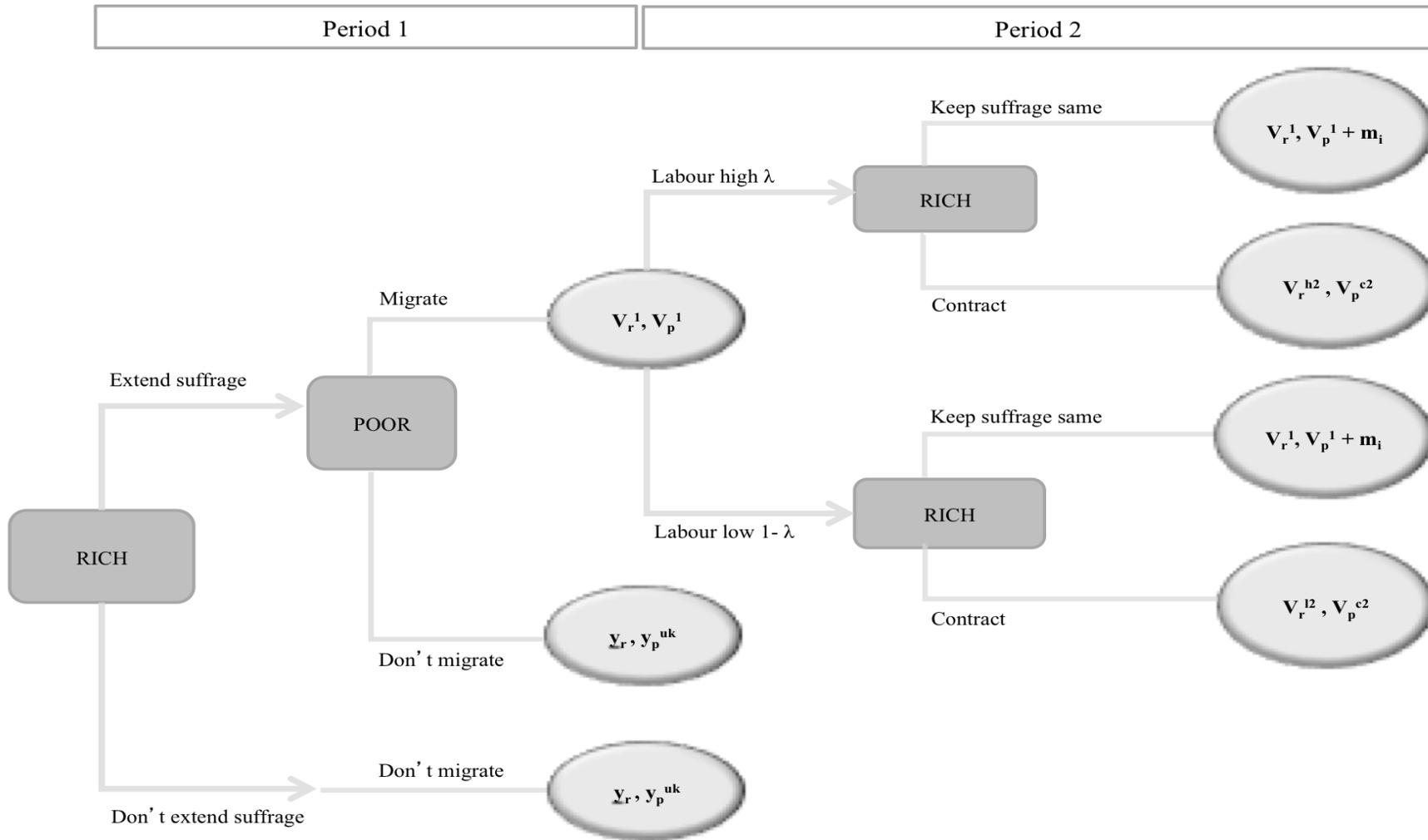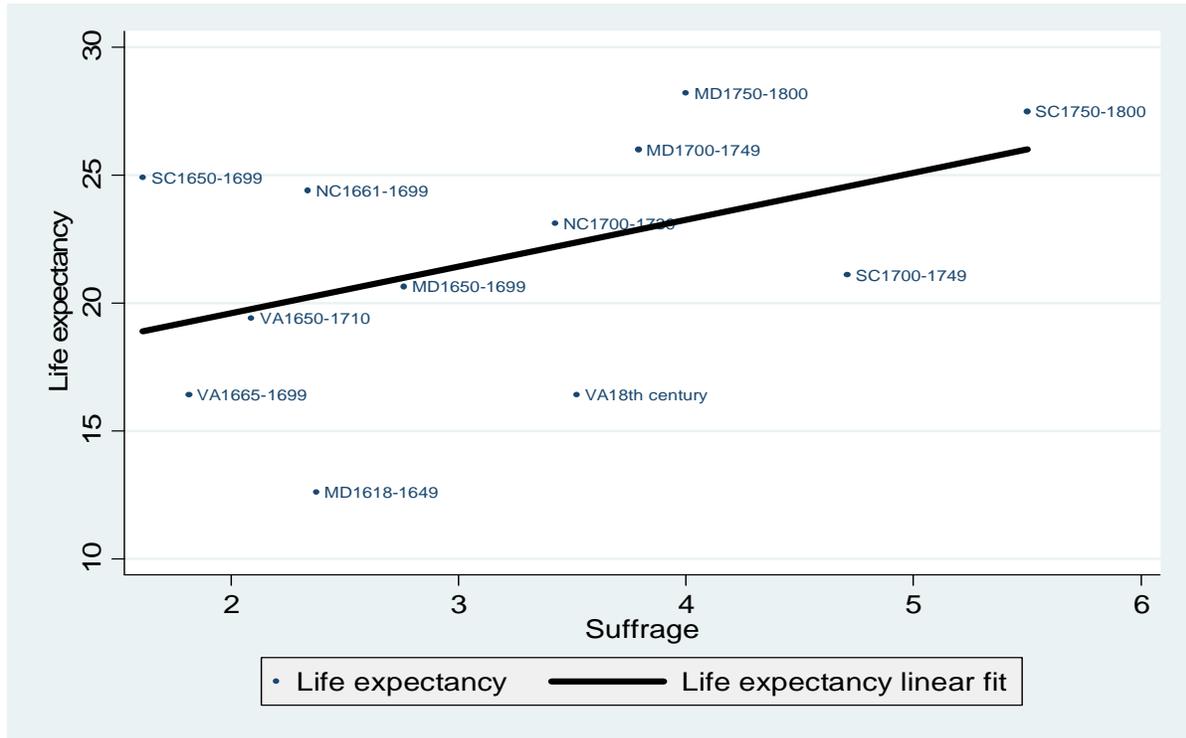
Chart S1: The suffrage game

Period 1	Period 2

Keep suffrage same
$V_r^1, V_p^1 + m_i$

Labour high λ

RICH

Contract
V_r^{h2}, V_p^{c2}

Migrate
V_r^1, V_p^1

Extend suffrage

POOR

Keep suffrage same
$V_r^1, V_p^1 + m_i$

Labour low $1-\lambda$

RICH

RICH

Contract
V_r^{l2}, V_p^{c2}

Don't migrate
$\underline{y}_r, y_p^{uk}$

Don't migrate
$\underline{y}_r, y_p^{uk}$

Don't extend suffrage

22

Chart S2: Life expectancy and suffrage

Sources: Life expectancy: Galenson (1996), p. 182 and Purvis (1999), p. 176; Suffrage data: McKinley (1905).

Notes: This graph shows the relationship between Suffrage and Life Expectancy using variation over time and across colonies. The solid line is obtained by a linear fit. Life expectancy is male life expectancy at age 30. Suffrage is averaged over the entire period covered by each life expectancy data point. See the notes to Chart 1 and Appendix 1 on how Suffrage is constructed.

Charts S3A - C: Evolution of prices for tobacco, rice and wheat

Chart S3A: Price for tobacco – Virginia
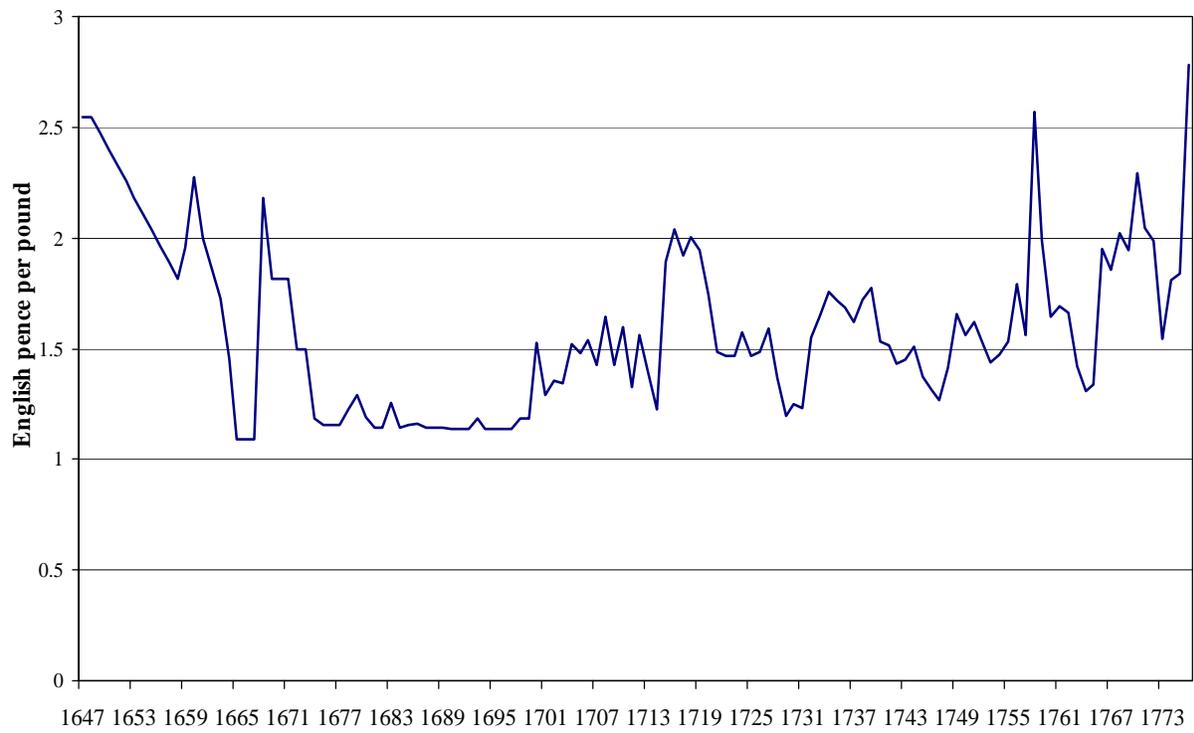
Chart S3B: Price for rice – Charleston, South Carolina
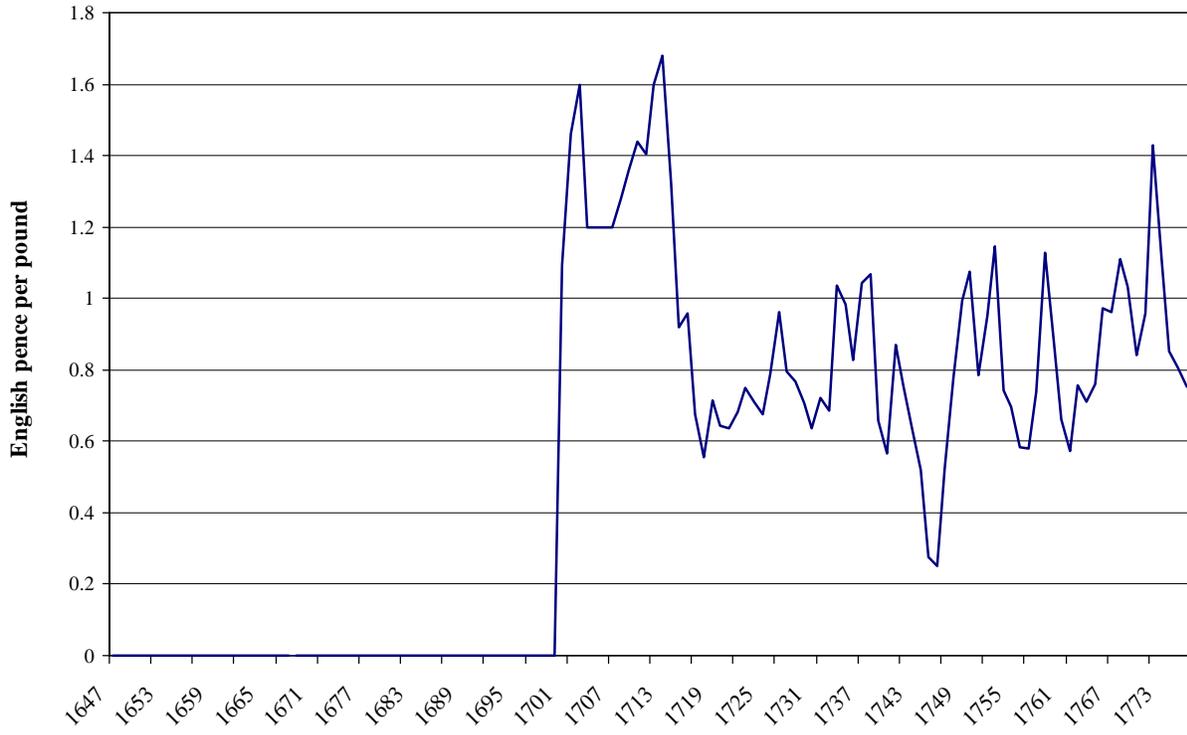
Chart S3C: Price for wheat – Talbot County, Maryland
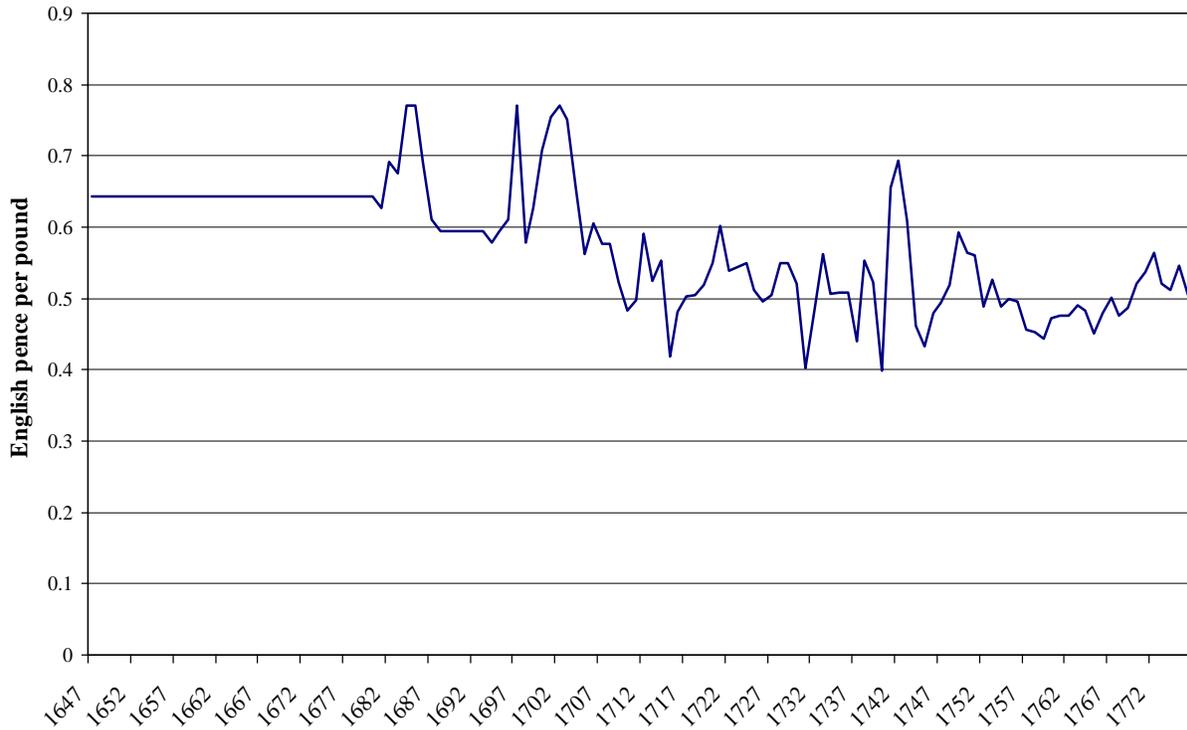
Sources: Rice and tobacco prices: Historical Statistics of the US Millennial Edition Online (2006); Wheat prices: Purvis 1999; exchange rate data: Historical Statistics of the US Millennial Edition Online (2006).

Notes: These graphs show the price movements for the three main export crops of colonial British America: tobacco, rice and wheat. I convert the colonial currencies in which these prices are measured to English pence. I use linear interpolation to fill in the missing values in the exchange rate data and assume that the earliest exchange rate data persisted backwards. Since rice production did not start until the late 1690s and there are no data on rice prices before 1701, I assume that the pre-1701 rice prices were 0.

Chart S4: Evolution of the ratio of Caribbean slave prices over the crop index
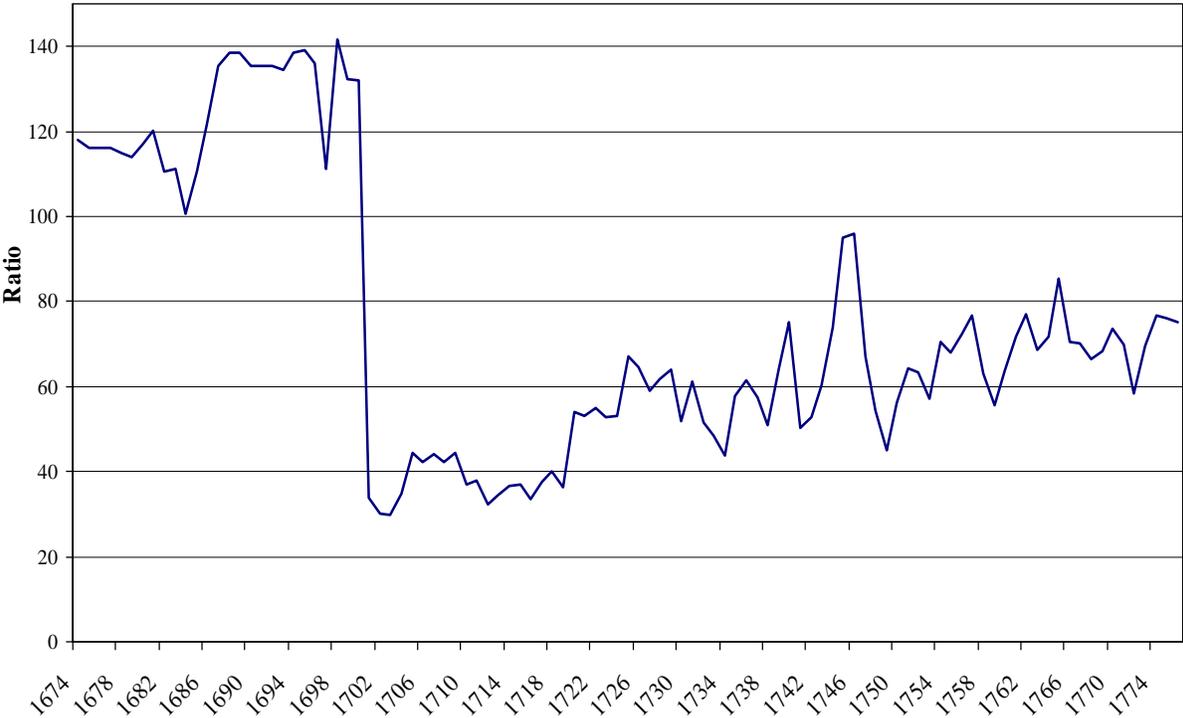
Sources: Caribbean slave prices: Eltis, Lewis and Richardson (2005); Crop Index: prices: Purvis (1999), Historical Statistics of the US Millennial Edition Online (2006); weights: National Agricultural Statistics Service, United States Department of Agriculture.

Notes: This graph shows the evolution of the ratio of Caribbean slave prices over the Crop Index. Caribbean slave prices are measured in British pounds sterling per slave, while the Crop Index is measured in British pence per pound.
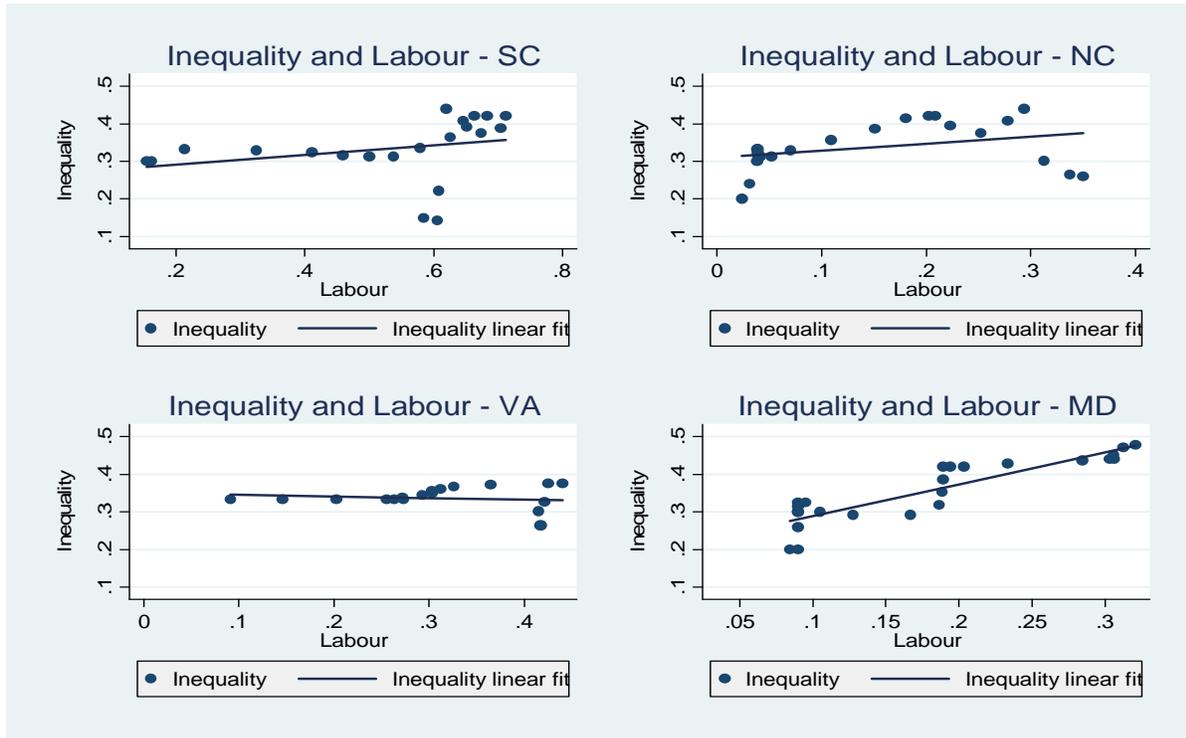
Chart S5: Inequality and Labour – South
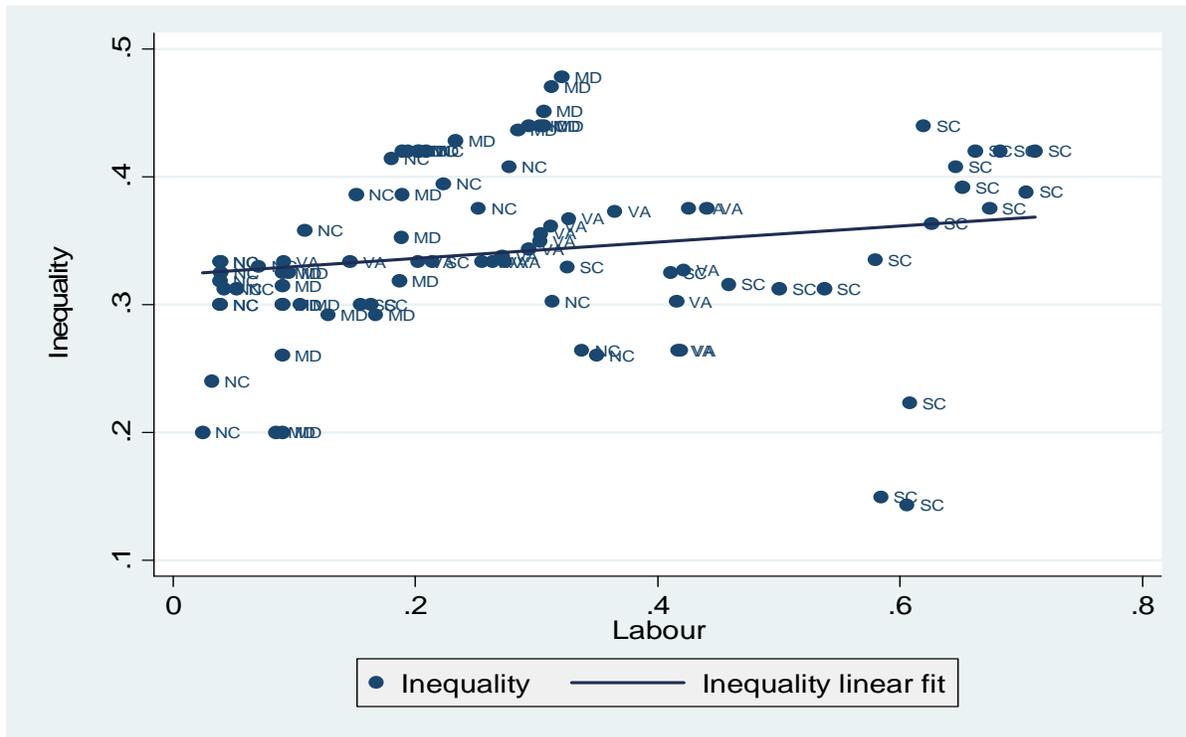
Sources: Inequality: Kulikoff (1986, 2000), Main (1965) and Nash (1979); Labour: Historical Statistics of the US Millennial Edition Online (2006).

Notes: This graph shows the unconditional correlation between Inequality and Labour among the Southern colonies. See the notes to Chart 1 about the colonies included in the Northern and Southern regions. Inequality is measured as percentage landless, where a higher value of percentage landless implies higher inequality. Labour is measured as the percentage of a colony's black population. The data set is obtained by taking every fifth observation for each colony. Each dot represents an observation for a particular colony and year. The solid line is obtained through a linear fit. See Appendices 1 and S1 for more details on the construction of the variables.
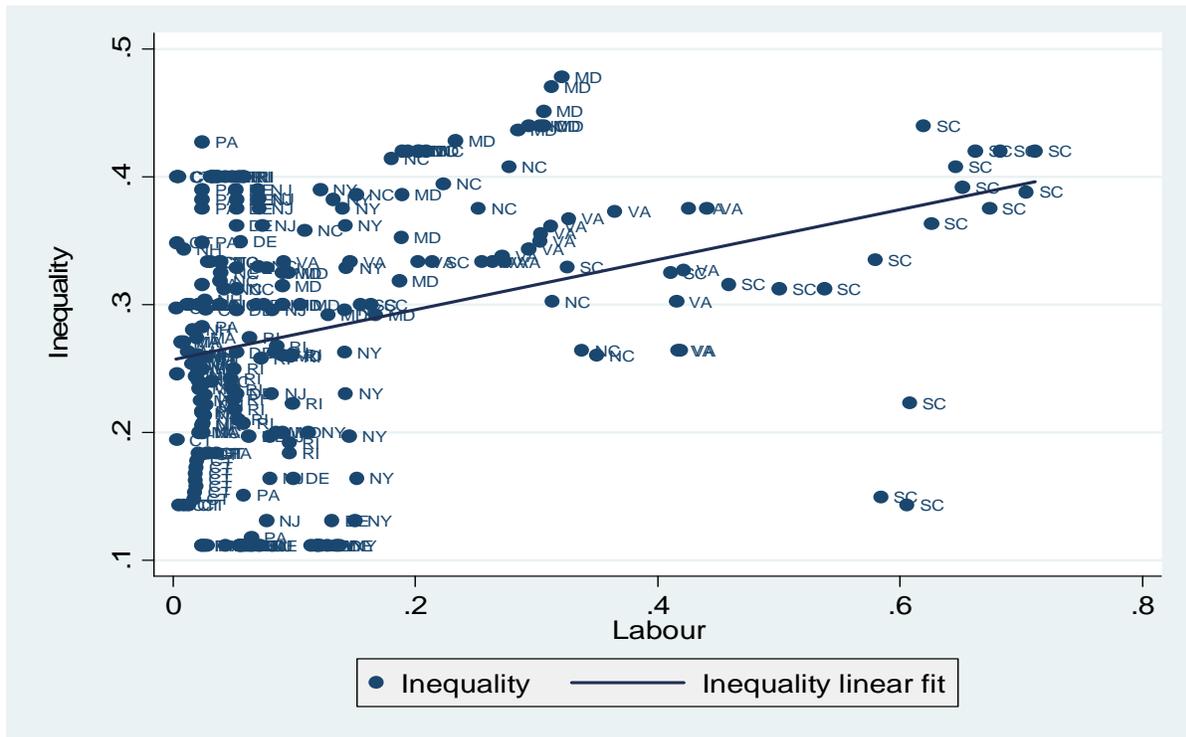
Chart S6: Inequality and Labour – South, by colony

Sources: Inequality: Kulikoff (1986, 2000), Main (1965) and Nash (1979); Labour: Historical Statistics of the US Millennial Edition Online (2006).

Notes: This graph shows the unconditional correlation between Inequality and Labour in North Carolina, South Carolina, Virginia and Maryland. Inequality is measured as percentage landless, where a higher value of percentage landless implies higher inequality. Labour is measured as the percentage of a colony's black population. The data set is obtained by taking every fifth observation for each colony. Each dot represents an observation for a particular colony and year. The solid line is obtained through a linear fit. See Appendices 1 and S1 for more details on the construction of the variables.

Chart S7: Inequality and Labour – North and South

Sources: Inequality: Kulikoff (1986, 2000), Main (1965) and Nash (1979); Labour: Historical Statistics of the US Millennial Edition Online (2006).

Notes: This graph shows the unconditional correlation between Inequality and Labour in the entire sample. See the notes to Chart 1 about the colonies included in the Northern and Southern regions. Inequality is measured as percentage landless, where a higher value of percentage landless implies higher inequality. Labour is measured as the percentage of a colony's black population. The data set is obtained by taking every fifth observation for each colony. Each dot represents an observation for a particular colony and year. The solid line is obtained through a linear fit. See Appendices 1 and S1 for more details on the construction of the variables.

Additional bibliography- supplementary materials

L. Bailey (1907), *Cyclopedia of American Agriculture: a Popular Survey of Agricultural Conditions, Practices and Ideals in the United States and Canada, Volume 1*, The Macmillan Company.

S. Engerman, R. Sutch and G. Wright (2006), "Slavery", in *Historical Statistics of the United States, Millennial Edition Online*.

Appendix S1: Additional data sources and description		
VARIABLE	**DESCRIPTION**	**SOURCE**
Dependent variable: Suffrage (obtained through principal component analysis)	Because of the small number of total observations (fewer than 20), in this analysis I exclude the restrictions for particular status, formal patenting of lands, felon and family. I perform orthogonal varimax rotation, and only keep the factor with the biggest eigenvalue. This factor explains 0.26 of the variance, and has high and positive loadings on the restrictions for freeman, tax, residency, religion and race, and negative but not as strong loadings on the remaining restrictions: oath taking, property or income, freeholder, householder, minimum freeholding or property, and being a good person. Very similar results were obtained using factor analysis with principal component factors instead of principal component analysis.	McKinley (1905)
Alternative measure for Labour: Caribbean Slave Prices/Crop Index	This is the ratio of Caribbean slave prices over the Crop Index variable. Caribbean slave prices are measured in pounds sterling per slave, while the Crop Index is measured in pence per pound.	Caribbean slave prices, 1674-1775: Eltis, Lewis and Richardson (2005); Crop Index – see Appendix 1.
Alternative measure for Labour: Binary Indicator	This variable equals 1 for the South after 1700 and 0 for the South after 1700 and for the North.	
Exports	I use data on the total value of exports to Britain during the colonial period in pounds sterling (constant value) divided by colonial population. The data set consists of annual data on exports from 1693-1791 for New England, New York, Pennsylvania, Virginia and Maryland, the Carolinas and Georgia. Data for 1696 were interpolated by averaging the data in 1695 and 1697. There are also no data for Delaware and New Jersey. Because trade flows dropped dramatically in 1776, export data for 1775 were used instead. To obtain exports data series by colony, the aggregated export numbers were broken up by assigning a weight for each colony based on the length of its tidal shoreline (from the Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service). As a robustness check, another set of weights were also assigned based on the distance from a colony's major port to London, England. The list of ports was obtained from Northeast Export.com, and the calculations of the distance to England (assuming air distance) were done using the distance calculator available at http://www.infoplease.com/atlas/calculate-distance.html. See the description for population density below for information on the data for colonial population. Regressions using the second set of weights produced very similar results and are available upon request.	Historical Census Browser. Retrieved 1 Aug, 2008, from the University of Virginia, Geospatial and Statistical Data Center: http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html. Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service: http://www.noaa.gov/ List of North American ports. http://en.wikipedia.org/wiki/List_of_North_American_ports Infoplease.com. Distance Calculator. http://www.infoplease.com/atla

Population density	Data on total population are from the US Census Bureau, for every ten years and for each colony. Data on the size of each colony are from Purvis (1999), 19 and from Purvis (1994), 243-244. Note that I used the Maine settled area by 1800 instead of the area of the modern state. I take the average value of each two ten-year periods to fill in missing values for observations for years ending with "five". I fill in the rest of the missing observations by rounding their year to the nearest 10th and using the already available data.	Historical Census Browser. Retrieved 1 Aug, 2008, from the University of Virginia, Geospatial and Statistical Data Center: http://fisher.lib.virginia.edu/collections/stats/histcensus/index.html Purvis (1999), 19 Purvis (1994), 243-244
Urbanisation	This is data on each colony's total urban population from Purvis (1994, 1999), normally in 10 or 5-year intervals, divided by total population. See the description for population density for information on the data for colonial population. To fill in missing values, I rounded the observation year to the nearest 10th or fifth digit.	Purvis (1994), 220 - 227 Purvis (1999), 253
Ethnic fractionalisation and religious fractionalisation	These are HHI-type indices of ethnic and religious fractionalisation, which were calculated from colony-level data on the ancestral origins of the white population in 1790 and on the distribution of different types of churches in 1750, respectively. Ancestral origins include: English, Welsh, Scotch-Irish, Scottish, Irish, German, Dutch, French and Swedish; and the types of churches include: Anglican, Baptist, Congregational, Reformed, Lutheran, Presbyterian, Roman Catholic and Jewish.	Ethnicity data: Purvis (1994), 182 Church data: Purvis (1999), 181
Dummies for proprietary or charter colony	Dummies for when each colony was a proprietary or charter colony (royal colony is the omitted category in the regressions).	Purvis (1999), 188