



Munich Personal RePEc Archive

Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis

Nathan Nunn

University of British Columbia

May 2007

Online at <http://mpra.ub.uni-muenchen.de/4080/>
MPRA Paper No. 4080, posted 15. July 2007

SLAVERY, INEQUALITY, AND ECONOMIC DEVELOPMENT IN THE AMERICAS: AN EXAMINATION OF THE ENGERMAN-SOKOLOFF HYPOTHESIS

Nathan Nunn^{*†}

May 2007

Abstract

Recent research argues that among former New World colonies a nation's past dependence on slave labor was important for its subsequent economic development (Engerman and Sokoloff, 1997, 2002, 2006; Sokoloff and Engerman, 2000). These studies argue that specialization in plantation agriculture based on slave labor caused economic inequality, which concentrated power in the hands of a small elite, adversely affecting the development of domestic institutions needed for sustained economic growth. I test for these relationships looking both across former New World economies and across states and counties within the U.S. I find evidence that slave use is negatively correlated with subsequent economic development. However, I do not find evidence that this negative relationship is driven by large scale plantation slavery, or that the relationship works through slavery's effect on economic inequality.

^{*}Department of Economics, University of British Columbia, and the Canadian Institute for Advanced Research (CIFAR).

[†]I thank Jim Robinson, Elhanan Helpman, Daron Acemoglu, and seminar participants at the CIFAR Institutions, Organizations, and Growth Program Meeting for valuable comments. I also thank Maira Avila, Ken Jackson, and Wendy Bo Wu for excellent research assistance.

1 Introduction

In a series of papers (Engerman and Sokoloff, 1997, 2002, 2006; Sokoloff and Engerman, 2000), economic historians Stanley Engerman and Kenneth Sokoloff argue that the different development experiences of the countries in the Americas can be explained by initial differences in factor endowments, which resulted in differences in the use of slave labor. The authors argue that reliance on slavery resulted in extreme economic, which hampered the evolution of institutions and other factors necessary for sustained long term economic growth.

In this chapter, I empirically examine the validity of two parts of Engerman and Sokoloff's hypothesis: that large scale plantation slavery resulted in economic inequality, which resulted in slow economic development.¹ These relationships are summarized in the following diagram,

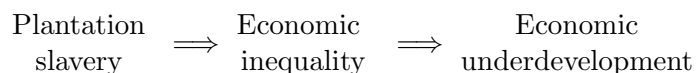


Diagram 1: Testing the channels of causality in Engerman and Sokoloff's hypothesis.

In first part of the chapter, I test for the reduced form relationship between large scale plantation slavery and economic underdevelopment. I do this by examining whether there is evidence that countries that relieve most heavily on slave use in the late 18th century to the early 19th century are poorer today. I test for this relationship across former New World economies, and across counties and states within the U.S. In both settings, I find a significant negative relationship between past slave use and subsequent economic development. I also examine whether large scale plantation slavery appears to have been particularly damaging for subsequent development. I do not find any evidence that large scale slavery was more detrimental for growth.

¹I do not examine the first component of their argument, that natural resources, such as soils suitable for plantation agriculture, were an important determinant of slave use in the colonies. The link between geography and slavery, across counties within the United States, has been examined by Lagerlöf (2005). He finds temperature, elevation, and precipitation to all be important determinants of slave use.

Instead, the evidence suggests that all forms of slavery were detrimental, and that if any form of slavery was particularly detrimental it was actually small scale urban slavery.

In the final part of the chapter, I examine whether, consistent with Engerman and Sokoloff's hypotheses, the slavery-income relationship can be explained by slavery causing extreme economic inequality, which adversely affected economic growth. Looking within the U.S, I find that slavery in 1860 is positively correlated with land inequality in 1860. That is, I find support for the first relationship in the diagram. However, I do not find that land inequality affected subsequent economic development, which is the second relationship in the diagram. In addition, I do not find that the link between slavery and inequality explains the relationship between slavery and subsequent economic development.

Overall, the results of this chapter support Engerman and Sokoloff's basic assertion that slavery was detrimental for economic development. However, the data do not show that large scale plantation slavery was particularly detrimental for development. As well, it does not appear that slavery's adverse effect on subsequent economic performance is because of its impact on initial economic inequality.

2 Looking within Former New World Countries

To construct measures of the prevalence of slave use in each New World country, I use historic population data from a variety of sources, most often population censuses.² As my measure of the prevalence of slavery I use the fraction of each country's total population in slavery in 1750. It is important to note that I am not using the proportion of the population that is of African descent. Included in the category of slaves are enslaved Africans and Natives Americans, while free Africans are not included. One could also construct estimates of the proportion of a population that was African, but this is a much less precise measure of the measure of interest, the use of slave labor.

As a measure of economic development I use the natural log of real per capita GDP in 2000. I choose this year because this maximizes the number of countries in the sample with income data. The sample of countries includes all currently independent former New World economies for which slave and free persons population data, and income data are available.

The relationship between income and the proportion of the population in slavery in 1750 is shown in figure 1. As shown, in the raw data one observes

²The sources of the data are described in the appendix.

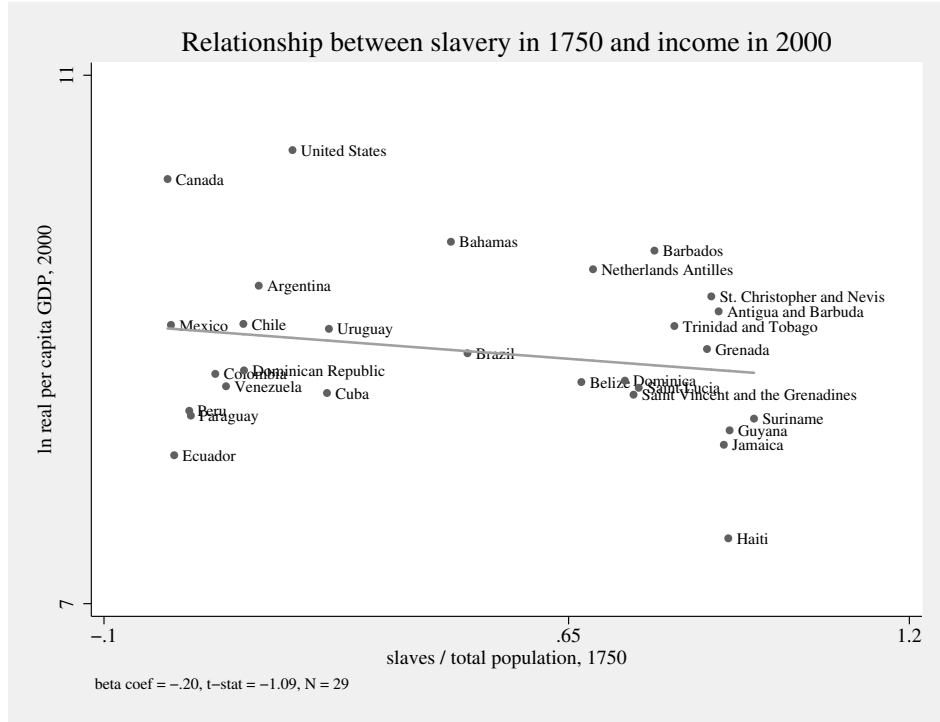


Figure 1: Bivariate plot showing the relationship between the proportion of the population in slavery in 1750, S_i/L_i , and the natural log of real per capita GDP in 2000, $\ln y_i$.

a weak negative relationship between past slave use and current income.

I further examine this relationship by also controlling for other potentially important country characteristics. I estimate the following equation,

$$\ln y_i = \alpha + \beta_S S_i/L_i + \gamma L_i/A_i + \mathbf{I}'\delta + \varepsilon_i \quad (1)$$

where y_i is country i 's real per capita GDP in 2000, S_i/L_i is the proportion of slaves in the total population in 1750, L_i/A_i is the population density in 1750, and \mathbf{I} denote colonizer fixed effects for former French, British, Spanish, Portuguese, and Dutch colonies.

The measure of population density is meant to capture the prosperity of a country in 1750, which in turn is determined by a host of factors such as climate, soil quality, natural openness, distance to international markets, and other determinants of the overall desirability of a colony at the time.

Part of Engerman and Sokoloff's argument is that in Spanish colonies,

Table 1: Slavery and income across former New World economies.

	(1)	(2)	Omit USA, CAN (3)	Omit USA, CAN, HTI (4)
Fraction slaves: S_i/L_i	-.67* (.35)	-2.63*** (.42)	-1.43* (.74)	-1.43* (.74)
Population density: L_i/A_i	.60* (.30)	.61*** (.21)	.59*** (.20)	.59*** (.21)
Colonizer fixed effects	No	Yes	Yes	Yes
R^2	.17	.66	.54	.37
Number of observations	29	29	27	26

Notes: The table reports OLS estimates of equation (1). Coefficients are reported with standard errors in brackets. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. The dependent variables is the natural log of real per capita GDP in 2000, $\ln y_i$. ‘Fraction slaves, S_i/L_i ’ is the number of slaves in the population divided by the total population. The colonizer fixed effects are for Portugal, England, France, Spain, and the Netherlands.

such as Peru and Mexico, large native populations survived European contact, and as a result the labor of Native Americans was used instead of slaves. The end result was large-scale estates and mines and extreme inequality. For this reason, I also include a fixed effect for Spanish colonies, as well as fixed effects for the nationalities of the other colonizers. The fixed effects also control for other differences in the colonial strategies of the colonizers.

OLS estimates of equation (1) are reported in table 1. The first column reports the estimates of (1) without colonizer fixed effects, while the second column reports the fully specified estimating equation with colonizer fixed effects. In both specifications the estimated coefficient for the fraction of the population in slavery S_i/L_i is negative and statistically significant.

The partial correlation plot for S_i/L_i from (1) (column 2) is shown in figure 2. Although there are no observations that appear to be biasing the results, Canada and the United States appear to be important observations. One may be concerned that the estimates may simply be reflecting differences between Canada and the United States, and the other less developed New World economies. If so, these differences may reflect other differences between the two groups, such as climate, natural resources, or the extent of European settlement.

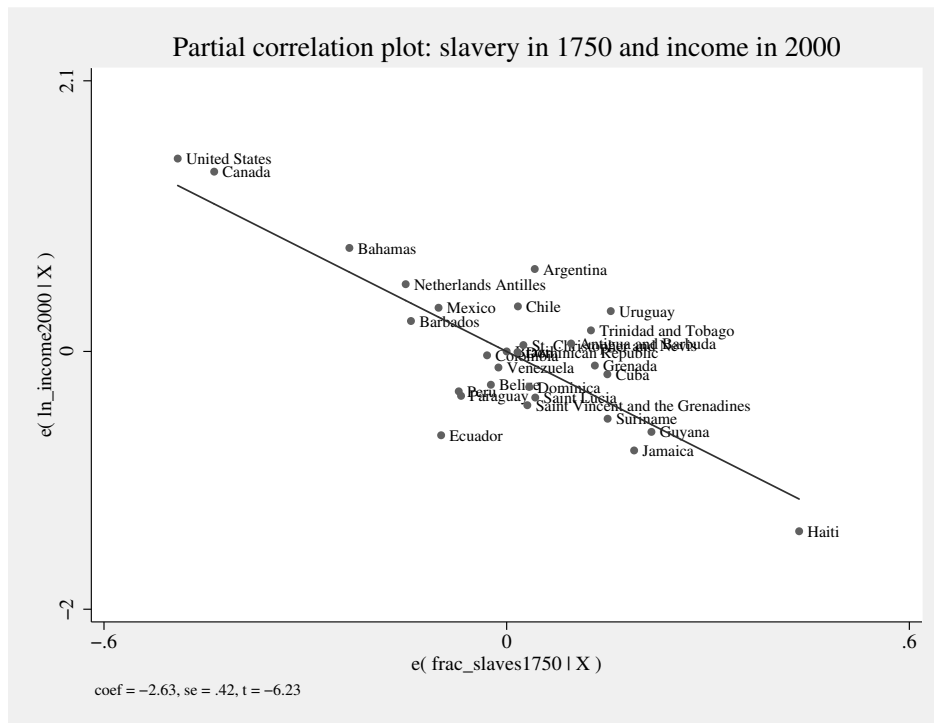


Figure 2: Partial correlation plot showing the relationships between the proportion of slaves in the population in 1750 and the natural log of real per capita GDP in 2000, $\ln y_i$.

In the third column of table 1, I re-estimate (1) after omitting Canada and the United States from the sample. As shown, the magnitude of the estimated coefficient decreases, but it remains statistically significant. Even ignoring Canada and the United States, one still observes a negative relationship between past slave use and subsequent economic development. This result is significant because the evidence presented in Engerman and Sokoloff (1997, 2002, 2006) and Sokoloff and Engerman (2000) generally relies on comparisons between Canada and the United States, and the other less developed countries in the Americas. The results here show that even looking within the later group one still observes a link between slavery and development. The final column also drops Haiti, which from figure 2 can be seen to also be a potentially influential observation. The results show that even after dropping the three countries from the sample, one still observes a significant negative relationship between slavery and subsequent income.

In the next section, I show that the relationship between slavery and economic development continues to hold when one considers an even more restricted, and more homogenous, sample of countries.

3 Looking within the British West Indies

In this section, I examine an even smaller sample of 12 countries that were formerly a part of the British West Indies. Although this is a very small sample, there are a number of benefits to examining this group of countries. First, the data are all from one source, British census records, all of which are recorded and summarized in Higman (1984). Because all data are from slave censuses that were conducted by the British government using the same procedures and administration, the data and information collected are quite reliable, and any biases or errors that may exist will generally be similar across all countries (Higman, 1984, pp. 6–15). Second, the sample of countries is homogenous in many dimensions. They are all small, former British colonies located in the Caribbean. As a result, many of the omitted factors that could potentially bias the estimates of interest, such as differences in culture, geography, or historical experience, are diminished by looking at this more homogenous sample.

The final benefit is that much more information is available for each country. Specifically, information on the size of plantations and on the use of slaves are available. This allows us to consider more deeply the hypotheses in Engerman and Sokoloff's work. To this point, we have examined the relationship between slave use and economic development, finding that,

consistent with their analysis, past slave use is associated with current underdevelopment. With the data from Higman we can begin to examine the potential channels behind this relationship. Because the hypothesized channel in Engerman and Sokoloff is through economic inequality, the authors focus almost exclusively on the adverse effect of slavery on large scale plantations. Their argument is that this form of slavery resulted in economic inequality, poor institutions, and economic stagnation.

Using Higman’s data on slave use and the size of slave holdings, one can examine whether the negative relationship between slave use is driven by large scale plantation slavery rather than other forms of slavery. I do this by allowing the slavery income relationship to differ depending on the manner in which the slaves were used. I first divide the total number of slaves in each society into three groups: urban domestic slaves (S_i^U), industry slaves (S_i^I), and plantations slaves (S_i^P). Industry slaves are slaves working in the livestock, salt, timber, fishing, or shipping industries. Plantation slaves are slaves working on sugar plantations, coffee plantations, cotton plantations, or in any other forms of agriculture.

In my estimating equation, I do not restrict the three measures of slavery to have the same effects on income. I allow each form of slavery to enter the estimating equation separately:

$$\ln y_i = \alpha + \beta_U S_i^U / L_i + \beta_I S_i^I / L_i + \beta_P S_i^P / L_i + \gamma L_i / A_i + \varepsilon_i \quad (2)$$

This equation is a less restrictive version of (1). If we restrict all three coefficients to be equal, $\beta_U = \beta_I = \beta_P$, then (2) reduces to (1).³

The slavery data that I use are now from 1830 rather than 1750. Although the total number of slaves and free persons are available for both 1750 and 1830, the number of slaves disaggregated by slave use is only available for 1830. Because, by 1830 none of the countries in the sample had abolished slavery, the proportion of slaves in 1830 is a good approximation of the use of slaves in the years prior to this date. The correlation between the proportion of the population in slavery in 1750 and in 1830 within the sample of 12 British colonies is .74. As shown in the first two columns of table 2, estimates of (1) are similar whether 1750 data or 1830 data are used.

Estimates of (2) are reported in the third column of table 2. All three slavery variables enter with negative coefficients, and the coefficients for urban slaves and for plantation slaves are statistically significant. Although

³Note also that because all of the countries in the sample are former British colonies the equation does not have colonizer fixed effects.

these results confirm the previous negative relationship between slave use and economic development, the relative magnitudes of the coefficients do not support Engerman and Sokoloff's focus on the detrimental effects of large scale plantation agriculture. According to the relative magnitudes of the coefficients, it is not the use of slaves on large scale plantations that has the greatest negative impact on development. Instead, it is the use of domestic slaves that has the largest negative impact. The difference between the estimated slave coefficients is statistically significant at the 8 percent significance level.

The partial correlation plots for the four variables in the estimation equation are shown in figure 3. From the plots it is apparent that the urban and industry slavery variables are each influenced by an outlier: the Bahamas for the urban slavery variable and Jamaica for the industry slavery variable. The negative relationship appears strongest and most robust for the plantation slavery measure.

Table 2: Slavery and income within the British West Indies.

	1750 (1)	1830 (2)	1830 (3)	1830 (4)
<u>Fraction of population that are:</u>				
Slaves, S_i/L_i	-2.42*** (.74)	-2.24** (.93)		
Urban slaves, S_i^U/L_i			-11.71** (4.18)	
Industry slaves, S_i^I/L_i			-4.20 (2.58)	
Plantation slaves, S_i^P/L_i			-4.19*** (1.01)	
Slaves on holdings with:				
10 slaves or less				-20.91*** (3.82)
11 to 200 slaves				-5.32*** (.95)
201 slaves or more				-8.12*** (1.30)
Population density: L_i/A_i	.24*** (.06)	.21*** (.07)	.23*** (.06)	.36*** (.03)
F-test of equality (p-value)			.08	.00
R^2	.69	.55	.79	.96
Number of observations	12	12	12	11

Notes: The table reports OLS estimates of equations (3) and (4). The dependent variable is the natural log of real per capita GDP in 2000, $\ln y_i$. Coefficients are reported with standard errors in brackets. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. The null hypothesis of the reported F-test is the equality of the three coefficients for the slavery variables.

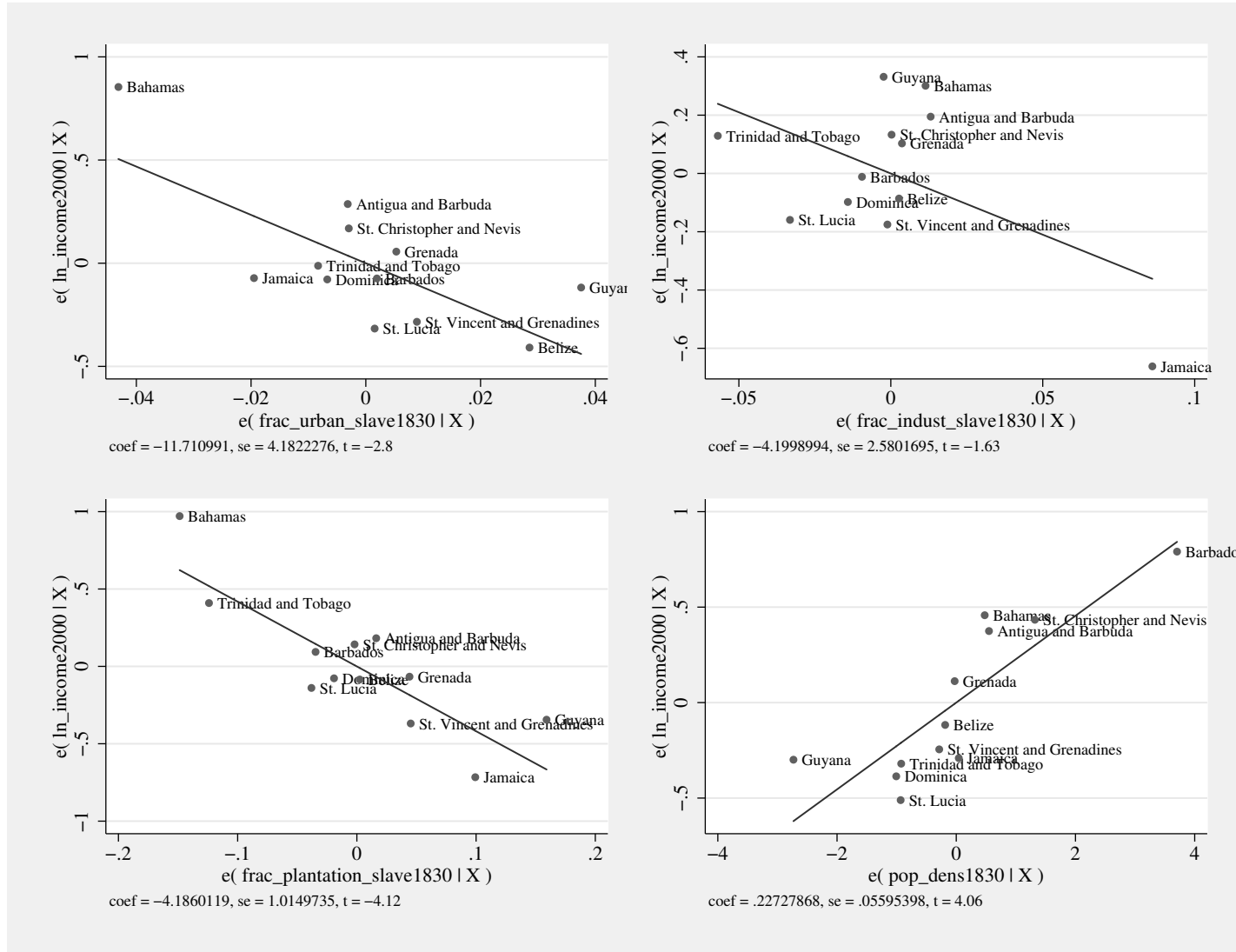


Figure 3: Partial correlation plots showing the relationships between urban slavery S_i^U/L_i , industry slavery S_i^I/L_i , plantation slavery S_i^P/L_i , population density L_i/A_i , and the natural log of real per capita GDP in 2000, $\ln y_i$.

I next consider whether the effect of slavery differs depending on the size of slave holdings. Higman provides data on the number of slaves on slave holdings with: (1) 10 slaves or less (2) 11 to 50 slaves (3) 51 to 100 slaves (4) 101 to 200 slaves and (5) 201 to 300 slaves or (6) 301 or more slaves. Because of the small number of observations available, I aggregate the holdings into three categories: (1) small scale holdings of 10 slaves or less, (2) medium scale holdings with 11 to 200 slaves, and (3) large scale holdings with 201 slaves or more. I then calculate of the proportion of the population that are slaves held on small scale holdings S_i^S/L_i , medium scale holdings S_i^M/L_i , and large scale holdings S_i^L/L_i .

Using these measures, I examine Engerman and Sokoloff's hypothesis that the detrimental impact of slavery arose because it was associated with large scale plantations, which resulted economic inequality. Because the size of slave holdings is an alternative indicator of the use of slaves on plantations, these results also provide a check for the results reported in table 2.

I use the following estimating equation to examine whether the effect of slavery on income differs by slave holding size:

$$\ln y_i = \alpha + \beta_S S_i^S/L_i + \beta_M S_i^M/L_i + \beta_L S_i^L/L_i + \gamma L_i/A_i + \varepsilon_i \quad (3)$$

Again, this equation is simply a more flexible version of 1.

OLS estimates of (3) are reported in column 4 of table 2.⁴ The results again support the notion that slave use was detrimental for economic development, but they do not support Engerman and Sokoloff's focus on the negative effects of large scale slave holdings. Contrary to the prediction that large scale slavery should have the largest impact on development, the estimates suggest that it is in fact small scale slavery that has the largest impact. The magnitude of the small scale coefficient is nearly 4 times the magnitude of the medium scale coefficient, and over twice the magnitude of the large scale coefficient. As well, these differences are statistically significant. The null hypothesis of the equality of the coefficients for S_i^S/L_i , S_i^M/L_i , and S_i^L/L_i is rejected at any standard significance level.

The partial correlation plots, reported in figure 3, show that the relationships between each of the three slavery variables and income appears robust. None of the relationships are driven by outlying observations.

⁴In this regression, the sample size is reduced to 11 countries because slave holding size data are unavailable for the Bahamas.

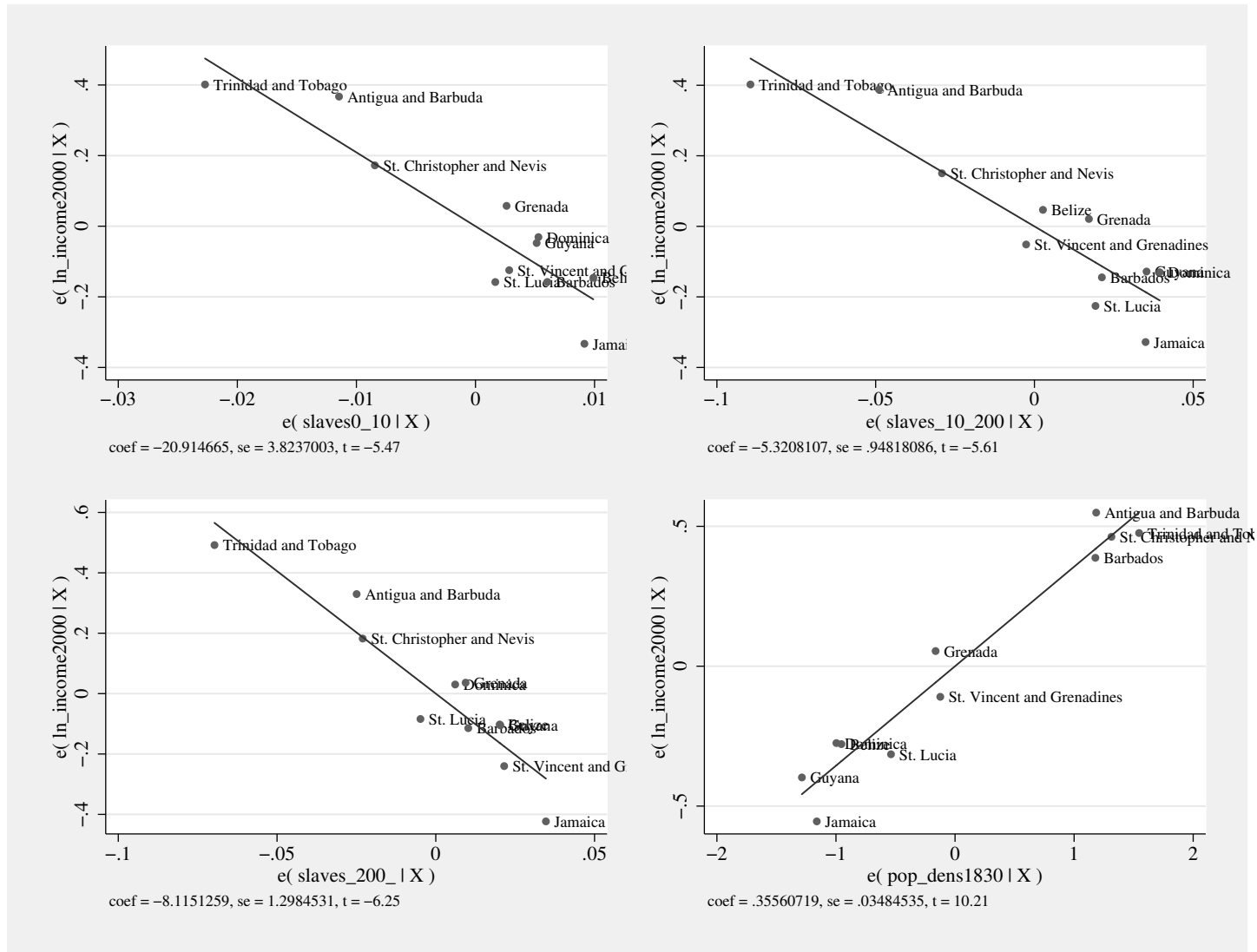


Figure 4: Partial correlation plots showing the relationships between small scale slavery S_i^S/L_i , medium scale slavery S_i^M/L_i , large-scale slavery S_i^L/L_i , population density L_i/A_i , and the natural log of real per capita GDP in 2000, $\ln y_i$.

Overall, these results confirm the findings in section 2 for the sample of New World countries. They provide added support for Sokoloff and Engerman’s hypothesis that slavery adversely affected subsequent economic development. However, they do not support their emphasis on the adverse effects of large scale plantation slavery. According to the estimates here, all forms of slavery appear detrimental for economic development. If anything, the evidence suggests that small scale urban slavery, not large scale plantation slavery, is the form of slavery that was the most harmful for economic growth.

4 Looking within the United States

I now turn to an additional source of variation, and look across counties and states within the U.S. Using information on the number of slaves and free persons in each county and state in the decades between 1790 and 1860, I again examine Engerman and Sokoloff’s most basic assertion that domestic slavery was detrimental for subsequent economic development. Population data for slaves and free persons are taken from the U.S. Decennial Censuses, while income data in 2000 are from the BEA’s *Regional Economic Accounts*.

The relationship between the proportion of the population in slavery in 1860 – the year for which data are available for the largest number of states – and the natural log of per capita income in 2000 is shown in figure 5. The figure shows a clear negative relationship between slave use and subsequent economic performance.

I explore this relationship further in table 3. Each column of the table reports the estimated relationship between slavery in each decade between 1790 and 1860 and per capita income in 2000, controlling for initial population density. The top panel of the table reports the relationship between the proportion of the population in slavery and per capita income across U.S. states. The number of observations begins at 17 in 1790 (the first column) and increases each decade to 37 in 1860 (the last column). The reason that the 1790 estimates include 17 states when only 13 states had joined the Union is that census data are available for 4 additional states. This is because in 1790 West Virginia and Kentucky were part of Virginia, while Maine was a part of Massachusetts. Therefore, data are available for these areas that later became independent states. As well, data are also available for Vermont which joined the Union a year later in 1791.⁵

⁵Similarly, in 1800 there are 18 observations even though only 16 states had joined the Union by this time. This is because of West Virginia and Maine.

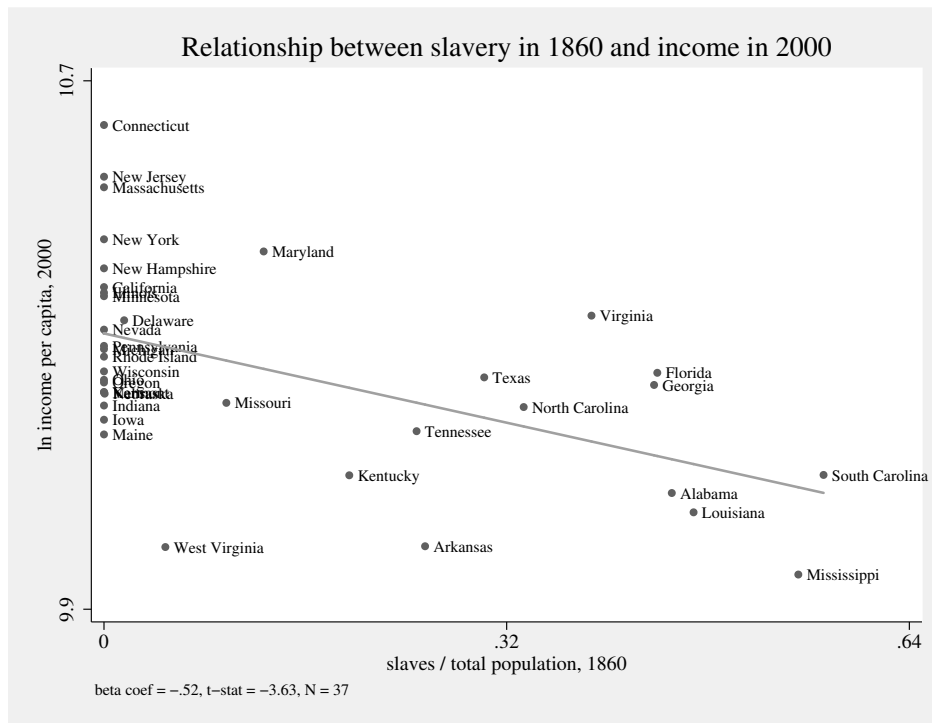


Figure 5: Bivariate plot showing the relationship between the proportion of the population in slavery 1860 and the natural log of per capita income in 2000, $\ln y_i$.

All of the estimated coefficients for the fraction of population in slavery S_i/L_i are negative. For the three decades prior to 1820 the coefficients are statistically insignificant, while for the five decades After 1810 the coefficients are statistically significant. The insignificance of the results for the first three decades is because three important slave states – Louisiana, Mississippi, and Alabama – did not join the Union until the decade after 1810. This can also be seen in figure 5. If one omits these three states, the negative relationship is weakened substantially.

The negative relationship between slave use and subsequent income levels across states is consistent with recent findings from Mitchener and McLean (2003) and Lagerlöf (2005). Mitchener and McLean (2003) estimate the relationship between slave use and subsequent labor productivity across U.S. states, and find a significant negative relationship between the fraction of the population in slavery in 1860 and average labor productivity in the decades after this date. Lagerlöf (2005), looking across U.S. counties, also documents a negative relationship between past slave use, measured in 1850, and subsequent per capita income measured in 1994.

The second panel of table 3 reports the same estimates looking across counties rather than states. As in the state level regressions, the coefficient estimates for S_i/L_i are negative. To be as conservative as possible, I allow for non-independence of counties within a state, and report standard errors clustered at the state level. This tends to at least double the reported standard errors. The coefficient estimates are negative and statistically significant for every year except 1810.

In the final panel of the table, I include state fixed effects. The slavery coefficients are now identified through variation within states only. The results show that once state fixed effects are controlled for, the estimated relationship between slavery and income becomes positive and statistically insignificant. This suggests that the strong negative cross-county relationship estimated in the second panel of the table is driven primarily by average differences across states, rather than differences between counties within a state.

Table 3: Slavery and income across counties and states within the U.S.

	1790	1800	1810	1820	1830	1840	1850	1860
Fraction slaves, S_i/L_i	-.13 (.24)	-.10 (.23)	-.11 (.20)	-.28* (.15)	-.29** (.14)	-.27** (.13)	-.34** (.13)	-.33*** (.11)
Population density, L_i/A_i	.52** (.20)	.57*** (.19)	.52*** (.17)	.46*** (.13)	.40*** (.11)	.33*** (.10)	.19** (.07)	.16*** (.05)
R^2	.38	.43	.44	.53	.53	.48	.42	.43
Number of observations	17	18	19	25	27	30	33	37
Fraction slaves, S_i/L_i	-.28** (.11)	-.21* (.12)	-.15 (.11)	-.17** (.10)	-.19** (.09)	-.24*** (.07)	-.23*** (.08)	-.22*** (.06)
Population density, L_i/A_i	.09*** (.01)	.06*** (.01)	.04*** (.007)	.03*** (.006)	.02** (.003)	.01*** (.002)	.007*** (.001)	.004*** (.001)
State fixed effects	No	No	No	No	No	No	No	No
R^2	.17	.13	.10	.09	.09	.09	.08	.07
Number of observations	283	400	521	739	964	1,273	1,588	2,014
Fraction slaves, S_i/L_i	.10 (.12)	.19 (.13)	.26* (.14)	.20 (.13)	.16 (.12)	.08 (.10)	.06 (.09)	.08 (.08)
Population density, L_i/A_i	.06** (.02)	.04*** (.002)	.03*** (.001)	.02*** (.001)	.01*** (.007)	.01*** (.004)	.006*** (.0003)	.004*** (.0003)
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.42	.42	.40	.39	.36	.36	.35	.32
Number of observations	283	400	521	739	964	1,273	1,588	2,014

Notes: The dependent variables is the natural log of real per capita income in 2000, $\ln y_i$. Coefficients are reported with standard errors in brackets. For the county level estimates the standard errors are clustered at the state level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. Population density L_i/A_i is measured in the appropriate base year.

The fact that the slavery-income relationship is driven by cross state variation, rather than within state variation, is completely consistent with the arguments of Engerman and Sokoloff. Their analysis emphasizes the effect of slave use on economic inequality, which adversely affected the formation of domestic institutions. If the institutions that are most important for economic development are at the national and state level rather than the county level, then we would not expect to only find a strong slavery-income relationship when looking across counties within a state. From this perspective, the insignificance of the variables in the bottom panel of the table provides evidence supporting Engerman and Sokoloff’s emphasis on the role played by domestic institutions.

The 1860 Census also reports the total number of slave holders that hold the following number of slaves: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10–14, 15–19, 20–29, 30–39, 40–49, 50–69, 70–99, 100–199, 200–299, 300–499, 500–999, and 1,000 and over. Because the census only reports information on the size holding of each slave holder and not of each slave (as in the Higman data), I can only calculate the number of slaves held in each size holding when the exact number of slaves per holder is given, which is only for holdings with less than 10 slaves. Therefore, I am able to separate small scale holdings (9 slaves or less) from medium/large scale holdings, but I am not able separate slaves held on medium holdings (10 to 199 slaves) from those held on large holdings (200 slaves or more).⁶

Using this data I construct two measures of slavery: the proportion of the population that are slaves held on small scale holdings S_i^S/L_i , and the proportion of the population that are slaves held on medium or large scale holdings S_i^{ML}/L_i . My estimating equation is thus,

$$\ln y_i = \alpha + \beta_S S_i^S/L_i + \beta_{ML} S_i^{ML}/L_i + \gamma L_i/A_i + \varepsilon_i \quad (4)$$

Table 4 reports the estimates of (4). Column 1 reports state level estimates. The coefficients for S_i^S/L_i and S_i^{ML}/L_i are both negative, but statistically insignificant. Their insignificance likely results because of multicollinearity (the correlation between S_i^S/L_i and S_i^{ML}/L_i is .87). However, the point estimates of the coefficients indicate that, contrary to the Engerman-Sokoloff view, slavery on small scale holdings is actually more detrimental for subsequent economic performance than large scale slavery.

Although these findings do not support Engerman-Sokoloff’s focus on large scale plantation slavery, it is possible that the data are not sufficiently

⁶Note that because of these same data limitations, the definition of small scale is slightly different than in section 3. Here the definition of small scale is 9 slaves or less, while the definition in section 3 was 10 slaves or less.

rich to identify the more harmful effects of medium/large scale slavery relative to small scale slavery. County level data provide finer variation which may help to better identify the differential effects of slavery (at the county level the correlation between S_i^S/L_i and S_i^{ML}/L_i is .65). Columns 2 reports county level estimates without state fixed effects. Here, the estimated effect of slavery on income is negative for both variables, and again the point estimates suggest that small scale slavery is worse for subsequent economic development than large scale slavery. In this case though, only the coefficient for medium/large scale slavery is statistically significant. As well, the F-test cannot reject the null hypothesis that the two coefficient are equal.

The final column reports county level estimates with state fixed effects. Here, the estimated effect of small scale slavery on income is positive and statistically significant and the effect for medium/large scale slavery is negative and statistically significant. As well, the null hypothesis of equality of the two coefficients can be rejected at the one percent significance level. This provides support for Engerman-Sokoloff's emphasis on large scale slavery. However, the estimated positive effect is very difficult to interpret, as it runs contrary to their hypothesis that slavery was detrimental for subsequent economic development.

The results to this point show that, both across countries in the Americas and across states in the U.S., there is a negative relationship between past slave use and current economic development. However, in generally the results do not indicate that large scale plantation agriculture was particularly detrimental. All forms of slavery appear to have had similarly detrimental effects on economic development, and that if any form of slavery was more harmful it was actually small scale urban slavery, not large scale plantation slavery..

5 Testing specific channels of causality

Recall diagram 1. To this point, I have only examined the reduced form relationship between slavery and economic development, without testing for the specific channels hypothesized by Engerman and Sokoloff. In this section, using data on the distribution of land holdings from the 1860 U.S. Census, I examine Engerman and Sokoloff's argument that slavery was detrimental because of its effect on initial economic inequality.

The Census provides data on the number of farms that fall into each of the following seven size categories: (1) 9 acres or less, (2) 10 to 19 acres, (3) 20 to 49 acres, (4) 50 to 99 acres, (5) 100 to 499 acres, (6) 500 to 999 acres,

Table 4: Slavery and income within the United States.

	State level	County level regressions	
	regressions	(4)	(6)
	(2)		
<u>Fraction of the population that are slaves:</u>			
on holdings with 9 slaves or less, S_i^S/L_i	-.41 (.99)	-.24 (.25)	1.68*** (.26)
on holdings with 10 slaves or more, S_i^{ML}/L_i	-.31 (.26)	-.22*** (.06)	-.11** (.05)
Population density, L_i/A_i	.16*** (.05)	.004*** (.0006)	.004*** (.0002)
State fixed effects		No	Yes
F-test of equality (p-value)	.93	.94	.00
R^2	.43	.07	.36
Number of observations	37	2,014	2,014

Notes: The dependent variables is the natural log of real per capita income in 2000, $\ln y_i$. Coefficients are reported with standard errors in brackets. For the county level estimates the standard errors are clustered at the state level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. In column 1 the unit of observation is a U.S. state, and in columns 2 and 3 the unit of observations is a U.S. county.

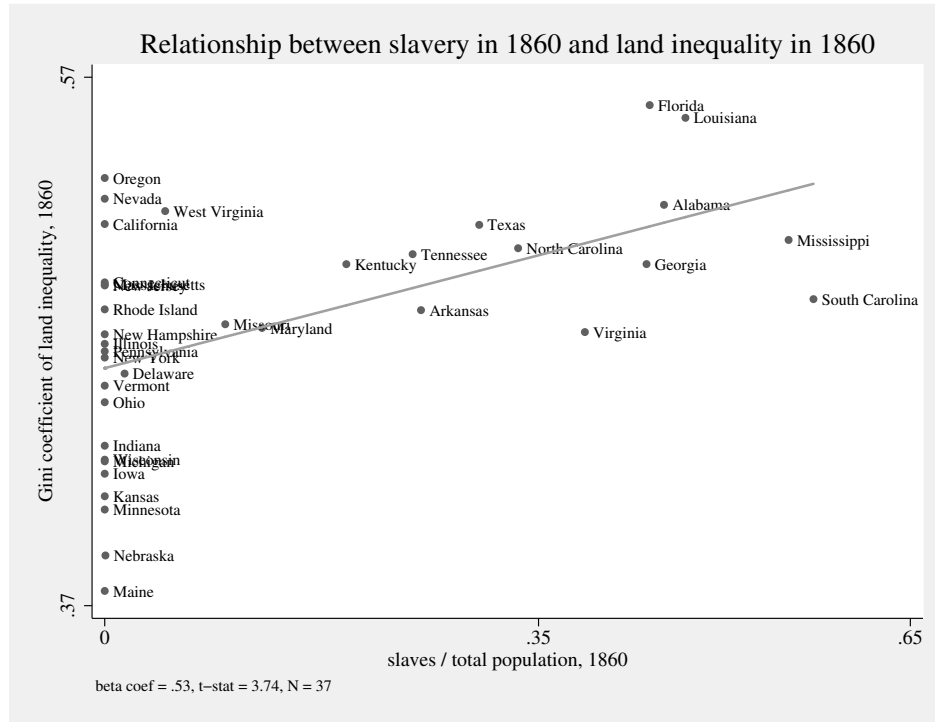


Figure 6: Bivariate plot showing the relationship between the proportion of the population in slavery in 1860 and the Gini coefficient of land inequality in 1860.

and (7) 1,000 acres or more. I use this information to construct, for each county and state, the Gini coefficient of land inequality in 1860. Full details of the construction are provided in the appendix.

I first examine the unconditional relationship between the proportion of the population in slavery in 1860 and land inequality in 1860. Figure 6 shows this relationship across states. Consistent, with Engerman and Sokoloff, one observes a positive statistically significant relationship between slavery and inequality.

I examine this relationship further in table 5. Column 1 reports the bivariate relationship between slavery and inequality shown in figure 6. In column 2, I allow the slavery inequality relationship to differ depending on the size of slave holdings. As shown, both coefficients are statistically insignificant. Again, this is most likely a result of the collinearity between the two slavery measures. Looking at the point estimates, the estimated

coefficient for the small scale slavery variable is larger, not smaller, than the medium/large scale coefficient, although the F-test cannot reject the null hypothesis of equality of the two coefficients. These results suggest that across states slavery is correlated with land inequality. However, they do not indicate that this relationship is driven by slavery on medium and large scale holdings.

In columns 3 to 6, I examine the same relationships at the county level.⁷ Column 3 reports the bivariate relationship between slavery and land inequality, while column 5 reports the same relationship controlling for state fixed effects. The results show that there is a positive relationship between slavery and land inequality across counties, but that this is driven by differences across states, not differences within states. Once state fixed effects are included in the regression equation the coefficient for S_i/L_i becomes statistically insignificant.

In columns 4 and 6, I allow the slavery-land inequality relationship to be different for slaves held on small holdings and for those held on medium/large scale holdings. In both specifications, without or with state fixed effects, I find that the proportion of the population held in slavery on small scale holdings is positively correlated with land inequality, but the proportion held on medium/large scale holdings is uncorrelated with inequality. In both specifications, the null hypothesis that the two coefficients are equal can be rejected. This suggests that slavery may have resulted in economic inequality, but it is not because of the particularly detrimental effects of large scale plantation slavery. Instead, it appears that, if anything, small scale slavery may be driving the relationship.

⁷Because farm size data are unavailable at the county level for Nebraska and Nevada, there are now only 1,933 observations in the county level regressions.

Table 5: Slavery and land inequality within the United States.

	State level regressions		County level regressions			
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Fraction of the population that are slaves:</u>						
on any size holding, S_i/L_i	.12***		.09***		.04	
	(.03)		(.03)		(.04)	
on holdings 9 slaves or less, S_i^S/L_i		.34		.46***		.33**
		(.30)		(.11)		(.15)
on holdings with 10 slaves or more, S_i^{ML}/L_i		.07		.01		.01
		(.08)		(.03)		(.04)
F-test of equality (p-value)		.49		.00		.04
State fixed effects			No	No	Yes	Yes
R^2	.29	.30	.07	.11	.25	.26
Number of observations	37	37	1,933	1,933	1,933	1,933

Notes: The dependent variables is the Gini coefficient of land inequality in 1860. Coefficients are reported with standard errors in brackets. For the county level estimates the standard errors are clustered at the state level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. In columns 1 and 2 the unit of observation is a U.S. state, and in columns 3 to 6 the unit of observations is a U.S. county.

These results are consistent with slavery resulting in increased economic inequality, which is the first relationship in diagram 1. Because economic inequality is very persistent even in the long-run,⁸ one also observes a strong relationship between past slave use and current economic inequality. This is shown in figure 7, where a clear positive relationship between past slave use and current inequality across U.S. states is evident. A related finding has been documented by Lagerlöf (2005), who finds that, looking across U.S. counties, the current white-black income gap is positively correlated with the proportion of the population in slavery in 1850.

The second part of Engerman and Sokoloff's hypothesis is that the link between slavery and inequality explains the relationship between slavery and subsequent economic development. Past economic inequality resulted in poor domestic institutions that provided advantages to domestic elites rather than providing support for broad based economic growth. This is the second relationship in diagram 1. In table 6, I examine whether the data support this explanation. The table reports estimates at both the state and county levels. Because I have already established that there is no relationship between slavery and income at the county level when state fixed are controlled for, I do not report the results from this specification; I only report state level estimates and county level estimates without fixed effects.

In columns 1 and 4, I re-estimate the slavery-income relationship. Column 1 simply reproduces the 1860 state level estimates from table 3. The estimates reported in column 4 are slightly different from the county level estimates from table 3. This is because the sample is reduced from 2,014 to 1,933 because of missing farm size data for the counties of Nebraska and Nevada. As shown, one still finds a negative relationship between slavery and income among this slightly smaller group of counties.

I next estimate the relationship between initial land inequality and subsequent income. This is the second part of Engerman and Sokoloff's hypothesis: economic inequality hampered economic development. These results are reported in columns 2 and 5. In both specifications, the estimated coefficient for land inequality is statistically insignificant. Thus, there is no evidence that economic inequality in 1860 is associated with worse economic performance in 2000.

Next, I examine whether the slavery-income relationship can be accounted for by land inequality. I do this in columns 3 and 6, where I include both the land Gini and the fraction of slaves in the population as explanatory

⁸The cross-state correlation between Gini coefficient of land inequality in 1860 and the Gini coefficient of income inequality in 2000 is .65.

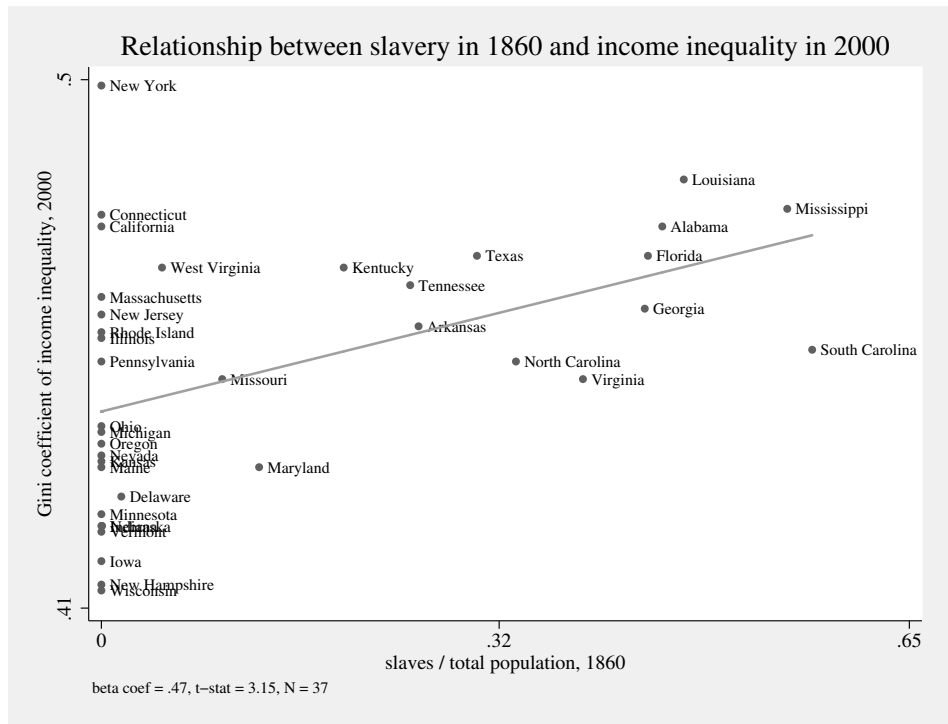


Figure 7: Bivariate plot showing the relationship between the proportion of the population in slavery in 1860 and the Gini coefficient of income inequality in 2000.

variables in the estimating equation. If slavery affects income only through its effect on economic inequality, then one would expect that controlling for economic inequality will significantly reduce the estimated relationship between slavery and income. The results show that this is not the case. At both the state and the county levels, including the land inequality measure actually increases the magnitude of the estimated coefficient for S_i/L_i , rather than decreasing it. At the state level, the estimated effect increases from $-.33$ to $-.39$, and at the county level, the effect increases from $-.23$ to $-.24$.

The finding that land inequality in 1860 is uncorrelated with income in 2000 is particularly interesting given that others have found evidence that early land inequality had adverse effects up until the early 20th century. Acemoglu *et al.* (2007) document a negative relationship between land inequality in 1860 and school enrollment in 1950. Similarly, Ramcharan and Erikson (2006) also finds a negative relationship between land inequality and per capita education expenditures in 1930. The results of table 6 suggest that by the end of the 20th century, the adverse effects of land inequality appear to have died out.

The results of this section are best summarized by returning to diagram 1. There is evidence of the first relationship in the diagram: slavery in 1860 is associated with greater land inequality at the same time. This was shown in table 5. Further, as a result of the persistence of economic inequality, there is also a strong positive relationship between slavery and current income inequality. However, I do not find evidence for the second relationship in diagram 1. The results of table 6 show that land inequality in 1860 is not correlated with income in 2000. This in turn suggests that the positive relationship between slavery and inequality does not explain the negative relationship between slavery and economic development. Instead, the data suggest that slavery had two distinct impacts. First, slavery resulted in lower long-term economic growth, and second, slavery resulted in greater initial inequality, which has persisted until today. These two effects appear to be unrelated. Contrary to Engerman and Sokoloff's hypothesis, slavery was not detrimental for economic development because it increased initial economic inequality.

Table 6: Slavery, land inequality, and income within the United States.

	State level regressions			County level regressions		
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction slaves, S_i/L_i	-.33*** (.16)		-.39*** (.13)	-.23*** (.07)		-.24*** (.07)
Gini coefficient of land inequality		-.46 (.51)	.45 (.55)		-.11 (.11)	.07 (.11)
Population density: L_i/A_i	.16*** (.05)	.21*** (.06)	.15*** (.05)	.004*** (.0006)	.004*** (.0006)	.004*** (.0006)
State fixed effects				No	No	No
R^2	.43	.30	.45	.08	.03	.08
Number of observations	37	37	37	1,933	1,933	1,933

Notes: The dependent variables is the natural log of real per capita income in 2000, $\ln y_i$. Coefficients are reported with standard errors in brackets. For the county level estimates the standard errors are clustered at the state level. ***, **, and * indicate significance at the 1, 5, and 10 percent levels. In columns 1 to 3 the unit of observation is a U.S. state, and in columns 4 to 6 the unit of observations is a U.S. county.

6 Conclusions

This chapter has examined the core predictions that arise from a series of influential papers by Stanley Engerman and Kenneth Sokoloff (Engerman and Sokoloff, 1997, 2002, 2006; Sokoloff and Engerman, 2000). I have found evidence consistent with their broad hypothesis that slavery was detrimental for subsequent economic development. There is a statistically significant negative relationship between past slave use and current economic development, looking either across countries within the Americas or across states within the U.S. However, contrary to their more specific arguments, I do not find that large scale plantation slavery was more harmful for growth than other forms of slavery. The evidence suggests that all forms of slavery were detrimental, and that if any form of slavery was more harmful it was actually small scale urban slavery, not large scale plantation slavery.

I have also examined whether the relationship between slavery and income can be explained by slavery's effect on initial economic inequality. Looking within the U.S. I found that slave use in 1860 is positively correlated with land inequality in the same year, but that this relationship does not explain slavery's adverse effect on subsequent economic development. Because of the persistence of inequality overtime, past slave use is also positively correlated with current income inequality. Thus, the data suggest that slavery had a long-term effect on both income and inequality, but that the two relationships are independent of one another. Contrary to Engerman and Sokoloff's hypothesis, slavery's effect on income does not appear to be the result of slavery increasing initial inequality which adversely affected long-term growth.

Although these results take us one step towards better understanding the long-term impacts of slave use in the Americas, an important question remains. If the relationship between past slave use and current income is not through the channel hypothesized by Engerman and Sokoloff, then what explains the relationship? One possibility, which is highlighted by Acemoglu *et al.*'s (2007) chapter in this book, is that what may have been important for long-term economic development was political inequality, not economic inequality. However, much remains to be done before we fully understand the mechanisms underlying the long-term effects of slavery in the Americas.

A Data Appendix

Data on country level real per capita GDP in 2000 are from World Bank (2006). For countries with missing income data, when possible converted income data from the Penn World Tables or Maddison (2003) were used. State and county level per capita income in 2000 are from the BEA's *Regional Economic Accounts*. The county level data are from Table CA1-3 located at www.bea.gov/regional/reis/, and the state level data are from Table SA1-3 at www.bea.gov/regional/spi/.

Population density is measured in hundreds of persons per square kilometer in the cross-country regressions, and hundreds of persons per square mile in the county and state level regressions. Country level land area data are from Harvard's *Center for International Development's Geography Database* located at www.ksg.harvard.edu/CID/ciddata/Geog/physfact_rev.dta. Land area for U.S. states and counties are from U.S. Bureau of the Census (2006).

The country level slave and free populations data used in section 2 are from a variety of sources. All data are from 1750 or the closest available year. Figures for Barbados, Saint Christopher and Nevis, Antigua and Barbuda, Jamaica, Cuba, Dominica, Saint Lucia, Saint Vincent and the Grenadines, Trinidad and Tobago, Grenada, Guyana, Belize, Bahamas, Haiti, Suriname, Netherlands Antilles, and the Dominican Republic are from Engerman and Higman (1997). All figures are for 1750. Data for Canada are from the *1784 Census of Canada*. Data for the United States are for 1774 and are from Jones (1980). Brazilian data are for 1798 and are taken from Simonsen (1978, pp. 54–57). Chilean data are from 1777 and are from Sater (1974). The figures for Colombia are for 1778 and are from McFarlane (1993). Data for Ecuador are for 1800 and are from Restrepo (1827, p. 14). Mexican data are for 1742 and are from Aguirre Beltran (1940, pp. 220–223). Peruvian data are for 1795 and are from Rugendas (1940). Data from Paraguay are for 1782 and are from Acevedo (1996, pp. 200–206). Venezuelan data are for 1800 and are taken from Figueroa (1983, p. 58). Data for Uruguay are for the city of Montevideo in 1800, and are taken from Williams (1987). Data for Argentina are for the city of Buenos Aires in 1810, and are from Rout Jr. (1976, pp. 91, 95) and Johnson *et al.* (1980).

Slave and free populations data for counties and states within the U.S. are from the 1790 to 1860 Decennial Censuses of the United States. The data have been digitized and can be accessed at: <http://fisher.lib.virginia.edu/collections/stats/histcensus/>. The data on the size of slave holdings, and the size of farms in 1860 are also from this source.

The Gini coefficient of income inequality for each state in 2000 is from

the U.S. Census Bureau. I approximate income inequality in 2000 using inequality in 1999, which is the closest year for which the inequality measures are available. The data were accessed from Table S4 available at: www.census.gov/hhes/www/income/histinc/state/state4.html

The Gini coefficient of land inequality is calculated using information about the size of each farm in the 1860 Census. The number of farms in each county is available for the following farm sizes: (1) 9 acres or less, (2) 10 to 19 acres, (3) 20 to 49 acres, (4) 50 to 99 acres, (5) 100 to 499 acres, (6) 500 to 999 acres, and (7) 1,000 acres or more. Because for each category I do not know the mean farm size, I use the median size of the category. For the category 1,000 acres or more, I use 1,000 acres. The Gini coefficients are calculated using the Stata program *ineqdec0* written by Stephen P. Jenkins. The formula for calculating the Gini coefficient is:

$$1 + (1/n) - \frac{2 \sum_{i=1}^n (n - i + 1) a_i}{n \sum_{i=1}^n a_i}$$

where n is the number of farms, a_i is farm size, and i denotes the rank, where farms are ranked in ascending order of a_i .

References

- Acemoglu, Daron, María Angélica Bautista, Pablo Querubín, and James A. Robinson, “Economic and Political Inequality in Development: The Case of Cundinamarca, Colombia,” (2007), mimeo, M.I.T.
- Acevedo, Edberto Oscar, *La Intendencia del Paraguay en el Virreinato del Río de la Plata* (Ediciones Ciudad Argentina, Buenos Aires, 1996).
- Aguirre Beltran, Gonzalo, *La Poblacion Negra de Mexico, 1519–1810* (Fondo de Cultura Economica, Mexico City, 1940).
- Engerman, Stanley L., and B. W. Higman, “The demographic structure of the Caribbean slave societies in the eighteenth and nineteenth centuries,” in Franklin W. Knight, ed., *General History of the Caribbean, Volume III: The slave societies of the Caribbean* (UNESCO Publishing, London, 1997), 45–104.
- Engerman, Stanley L., and Kenneth L. Sokoloff, “Factor Endowments, Institutions, and Differential Paths of Growth Among New World Economies: A View from Economic Historians of the United States,” in Stephen Harber, ed., *How Latin America Fell Behind* (Stanford University Press, Stanford, 1997), 260–304.

- , “Factor Endowments, Inequality, and Paths of Development Among New World Economies,” Working Paper 9259, National Bureau of Economic Research (2002).
- , “The Persistence of Poverty in the Americas: The Role of Institutions,” in Samuel Bowles, Steven N. Durlauf, and Karla Hoff, eds., *Poverty Traps* (Princeton University Press, Princeton, 2006), 43–78.
- Figuroa, Federico Brito, *La estructura económica de Venezuela colonial* (Universidad Central de Venezuela, Ediciones de la Biblioteca, Caracas, 1983).
- Higman, Barry W., *Slave Populations of the British Caribbean, 1807–1834* (The John Hopkins University Press, Baltimore, 1984).
- Johnson, Lyman L., Susan Migden Socolow, and Sibila Seibert, “Población y Espacio en el Buenos Aires del Siglo XVIII,” *Desarrollo Económico*, 20 (1980), 329–349.
- Jones, Alice Hanson, *Wealth of a Nation to Be* (Columbia University Press, New York, 1980).
- Lagerlöf, Nils-Petter, “Geography, Institutions and Growth: The United States as a Microcosm,” (2005), mimeo, York University.
- Maddison, Angus, *The World Economy: Historical Statistics* (Organisation for Economic Co-operation and Development, Paris, 2003).
- McFarlane, Anthony, *Colombia before Independence: Economy, Society, and Politics under Bourbon Rule* (Cambridge University Press, New York, 1993).
- Mitchener, Kris James, and Ian W. McLean, “The Productivity of U.S. States Since 1880,” *Journal of Economic Growth*, 8 (2003), 73–114.
- Ramcharan, Rodney, and Lennart Erikson, “Inequality and Redistribution Evidence From US Counties and States, 1890-1930,” (2006), mimeo, International Monetary Fund.
- Restrepo, José Manuel, *Historia de la Revolución de la República de Colombia en la América meridional, vol I* (Bensanzon, Paris, 1827).
- Rout Jr., Leslie B., *The African Experience in Spanish America* (Cambridge University Press, London, 1976).

- Rugendas, João Maurício, *Viagem Pitoresca através do Brasil* (Livraria Martins, São Paulo, 1940).
- Sater, William F., “The Black Experience in Chile,” in Robert Brent Toplin, ed., *Slavery and Race Relations in Latin America* (Greenwood Press, Westport, 1974), 13–50.
- Simonsen, Roberto Cochrane, *Historia Economica do Brasil: 1500/1820* (Companhia Editora Nacional, Sao Paulo, 1978).
- Sokoloff, Kenneth L., and Stanley L. Engerman, “History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World,” *Journal of Economic Perspectives*, 14 (2000), 217–232.
- U.S. Bureau of the Census, *County and City Data Book, 2000* (U.S. Department of Commerce, Bureau of the Census, Washington, D.C., 2006).
- Williams, John Hoyt, “Observations on Blacks and Bondage in Uruguay, 1800-1836,” *The Americas*, 43 (1987), 411–427.
- World Bank, *World Development Indicators* (World Bank, Washington, D.C., 2006).