JOURNAL OF ECONOMIC DEVELOPMENT 61
Volume 28, Number 2, December 2003

INCOME INEQUALITY AND POPULATION DENSITY 1500 AD: A CONNECTION

KEVIN SYLWESTER*

Southern Illinois University

Using a cross section of countries, this paper examines the association between regional population densities in 1500 AD and current income inequality. After controlling for other regional and historical factors, I find that past population density is negatively associated with income inequality today. Formerly high density regions are predicted to have lower income inequality. These findings support the view that higher density areas were better able to form more diverse and mobile societies that affected the long-run distribution of income.

Keywords: Income Inequality, Population Density
JEL classification: O15

1. INTRODUCTION

Many people prefer less skewed income distributions given the same average income. Past research has cited other reasons why income inequality might be undesirable. Alesina and Perotti (1996) claim that more unequal societies are more likely to suffer from political instability. Alesina and Rodrik (1994), Persson and Tabellini (1994) and Clarke (1995) report that countries with more income inequality have grown slower than otherwise similar countries. (See Barro (1999) and Benabou (1996) for surveys of past empirical and theoretical work linking income inequality to economic growth). These findings provide justification for decreasing the degree of income inequality in a society.

To help in this pursuit, identifying the determinants of income inequality would be beneficial. Why is it that some societies have significantly more unequal income distributions than others? Numerous studies have examined potential explanations for income inequality. (See Bruno *et al.* (1996), Chan (1989), and Weede and Tiefenbach (1981) for surveys) Some of these explanations center on current policies such as the support for international trade as in Bourguignon and Morrison (1990) or Chakrabarti (2000). However, Li *et al.* (1998) report that income inequality in many countries has

* I would like to thank an anonymous referee for suggestions improving the paper. All errors are my own.

remained stable over time and so the degree of income inequality might be more influenced by engrained historical characteristics than by changing policies (although there are exceptions such as Australia, France, New Zealand, Poland, Italy, and China where measures of income inequality show sizable changes in recent years). This does not necessarily imply that policies do not affect income distributions or that no set of policies can lead to less inequality, but only that other factors play larger roles in determining long-run distributions of income.[1]

This paper considers how one such historical factor might have affected income distributions today, namely population density. As shown below, there is no strong relation between current income inequality and current population density. More dense countries are not predicted to have different degrees of income inequality than do less dense ones. This finding is similar to one in Dehghan (1995) who does not find any strong association between overurbanization and income inequality in less developed countries. However, this does not imply that population density years ago might not have influenced subsequent income distributions. Higher population densities could have influenced societal outcomes by erecting or breaking down (or perhaps both) social hierarchies thereby affecting long-run distributions of income. A more elaborate discussion is provided in Section 2. The paper asks how population density centuries old might be associated with contemporary income inequality levels.

It shows that even after controlling for regional, geographic, historical, and contemporaneous factors, population densities 500 years ago are negatively correlated with current income inequality. This finding supports the view that more densely populated regions created more opportunities for social advancement and so led to more egalitarian societies. Another implication is that distributions of income are not just influenced by recently enacted policies but are potentially outcomes of historical processes centuries old. Not only might this often make altering income distributions difficult but one needs to consider long-run societal dynamics to better understand how one might lessen inequality.

The paper is organized as follows. Section 2 provides a theoretical context for how population density centuries ago might influence income inequality today. Related theories in the literature are also presented. Section 3 discusses the data and sets up the empirical methodology. Section 4 presents the results and a conclusion follows.

[1] A slightly different explanation is that policies do matter but that the likelihood of enactment of certain policies depends on historical characteristics not shared by all countries. Specific historical characteristics lead to specific policies that affect the distribution of income and it is difficult for countries with different histories to emulate this set of policies.

2. THEORETICAL CONTEXT

Some have argued that greater population in a given area of land creates societal inequities. Midlarsky (1982) and Midlarsky and Roberts (1985) claim that population growth leads to land scarcity and to greater inequities in land ownership between elite (i.e., ruling class) and poor farmers. North and Thomas (1973) report that conditions for feudal serfs deteriorated in medieval Europe as population grew but then improved after the Bubonic Plague reduced the number of people by a third.

Diamond (1999) argues that larger populations spurred more intensive farming techniques and better agricultural technology as the need for more food increased farming output. This view stems from Boserup (1965) who views population growth as spurring changes in farming techniques to meet the higher demand for food. Over time, a surplus was created which allowed some members of the group to pursue other endeavors including commerce, religion, military, and governance. Larger communities might have also been better able to advance technological frontiers thereby creating new products and opportunities. These more heterogeneous communities became more socially stratified as not all individuals were small farmers. This social stratification led to more income inequality. Diamond presents outcomes across the South Pacific to support his case. He claims that small islands, including atolls, having few people had lower degrees of social stratification than did bigger islands such as Hawaii which developed social hierarchies across a larger community.

But does this mean that those places that had high population densities centuries ago continue to have higher degrees of income inequality today? Previous researchers have argued that levels of income inequality remain persistent over time. Engermann and Sokoloff (1997) state that economies of scale in Latin America promoted plantation farming which led to high levels of income inequality. These more unequal societies then found it more difficult to erect democratic institutions. Presumably, democracies are at least as willing to lessen income inequality as less pluralistic forms of government. Sylwester (2002) finds that current income inequality is more likely to fall in democracies than in authoritarian governments. Thus, one argument is that high levels of income inequality made establishing democracies more difficult which tended to propagate the initial income distribution over time.

Theoretical models have also demonstrated the persistence of initial inequalities. Galor and Zeira (1993) and Perotti (1993) create models where indivisibilities in education and imperfect capital markets exacerbate initial differences in income. Durlauf (1996) and Benabou (1994) examine how differences in wealth lead to the formation of heterogeneous neighborhoods where higher income neighborhoods spend more on public education thereby increasing income differences over time. Although these models focus on education, they can be generalized to other cases where low income groups might be too poor to make investments or where agents self select into groups to maximize the presence of positive spillovers. High initial levels of inequality persist over time.

Given the above perspectives, one might argue that societies with large population densities in the past should continue to be more stratified today. They were initially more stratified and if income inequality persists over time, then they should continue to have higher levels of inequality today. In empirical work, past population density should be positively associated with current income inequality.

But there are also reasons why historically more heterogeneous societies should have lower inequality today. Crenshaw and Ameen (1993) claim that the creation of surplus food and a move by some individuals into nonagricultural activities led to greater specialization and an increase in economic opportunities. Societal complexity increased providing more ways to increase income than by acquiring land. The concomitant increase in mobility allowed more individuals to partly approach the income levels of the elite. Thus, high densities in rural areas should lead to less inequality. Using a cross-country sample, they found that the natural log of population density in agricultural areas in 1960 is negatively associated with income inequality in developing countries.

This paper links both views and allows for the relationship between population density and income inequality to be non-monotonic. Assume that higher densities do, indeed, initially create more stratified societies as increased agricultural output allows for the creation of social hierarchies where most farm although a few pursue other pursuits, including governance. These societies are initially more unequal than smaller ones as described in Diamond (1999). However, the degree of income inequality is not constant. As Crenshaw and Ameen (1993) argue, the greater heterogeneity creates additional opportunities and agricultural surpluses allow more people to engage in activities other than agriculture. Over time, this allows for greater social mobility and so income inequality falls. In addition, the creation of different interests can lead to greater competition across groups and provide checks and balances so that one group does not become too powerful. Without any one group becoming dominant, a dual society with a small elite and an impoverished mass of people is less likely to form. Provided that average income levels are also rising, this description provides another argument for an "inverted-U" relation between income levels and income inequality although the time period does not span years or decades but generations.

What would this theory imply about what should be found in the data? In an empirical specification, high dense regions centuries ago should have lower income inequality today. Most world regions have more people today than hundreds of years ago. Population densities everywhere are higher. But in places that had more people initially, these societies have had longer periods of time for social mobility to undo initial inequities. In essence, they began this non-monotonic process sooner and so are farther along on the downward branch. This assumes that there has been sufficient time for initial stratification to occur and for more dense regions to experience greater falls in inequality. Given that I consider population density from five centuries ago (and even earlier), I do not view this assumption as unwarranted.

There is not argued to be any association between income inequality and population density today. Given the mass migrations of the 19th and 20th centuries and the recent huge worldwide increase in population, population density today is not highly correlated (0.11) with that of 500 years ago. Moreover, it is not clear if there has been sufficient time for more recent changes in population to influence societies as described above. Therefore, this paper does not offer a prediction regarding population density today and income inequality. It would also be interesting to examine empirically if more dense societies did once have higher inequality as some have speculated using prevalent measures of income inequality such as the Gini coefficient. One could then compare findings and see if more dense societies once had more inequality but now less inequality. Obviously, the data is not available for such an examination.

Other researchers have also made connections between population densities in the past and institutional and economic outcomes today. These views focus specifically on how local populations influenced European colonial policy and their consequent effects. Acemoglu *et al.* (2002) argue that more densely populated regions in European colonies gave the ruling Europeans more incentive to create forced labor economies or to exact great tribute since the tax base was large. They further argue that these policies led to extractive institutions which continued to exist after independence. They view this as an explanation as to why high income regions before European colonization are relatively poor today. Although they do not comment on how this path might have affected long-run distributions of income, it presumably led to greater income inequality since these extractive institutions benefited the few at the expense of many.

Past population density might also have determined patterns of European settlement as settlers (unlike colonial administrators or governors) were more likely to migrate to regions where indigenous populations were small. Without a labor force to utilize, economic incentives pushed for small scale farming thereby creating the beginnings of a middle class and more equal distributions of income. The early presence of small scale farming instead of plantations might help to explain why Costa Rica and Columbia have relatively low income inequality in Latin America (see Nugent and Robinson (2000)). Nevertheless, this explanation is mitigated to the extent that a labor (slave) force was imported from abroad. If settlers where labor was scarce imported slaves from abroad, then this might have created the seeds for a very unequal society centuries later. Examples such as Brazil quickly come to mind.

Given these differing circumstances for regions outside of Europe, I will also examine them separately to determine the extent previous results are robust. However, it is not clear how the association between past population density and current income inequality should change in this restricted sample. According to Acemoglu *et al.* (2002), Europeans were likely to exploit large indigenous populations which could have led to more inequality over time. On the other hand, slaves were more likely to be imported where labor was scarce.

3. DATA AND EMPIRICAL SPECIFICATION

The Gini coefficient is used as it is the most widely available measure of income inequality. Inequality data comes from Deininger and Squire (1996) [DS].[2] Clarke (1995) reports that the Gini coefficient is strongly correlated with other measures of income inequality. Data is from 1990 but comes from the nearest year to 1990 if otherwise not available. Although it would be ideal to use data from the same year, Li *et al*. (1998) find that income inequality remains stable over time for many countries and so using data from different years should not greatly distort the findings. All of the Gini coefficients come from income data and not expenditure data since DS find that differences between the two can be large. It also comes from surveys that are national in scope (versus strictly urban or strictly rural). Nevertheless, the data might still differ in other ways (individual versus household, gross of taxes versus net). DS do not find great differences in inequality measures across these other distinctions. Unless stated otherwise, the remaining data comes from Englebert (2000). The data is available from the author upon request.

The empirical specification takes the following functional form:

$$\text{GINI} = a + B^{*}X + c^{*}\text{DEN1500} + e. \tag{1}$$

GINI is the Gini coefficient and X is a matrix of controls described below. DEN1500 is from McEvedy and Jones (1978) and denotes population density in 1500. I consider population density 500 years ago since this was at the beginning of European colonization of many world regions and so is largely unaffected by this later migration. Robustness checks will consider population density of other years. If data for a country is not given, I use regional data. The focus is on the coefficient c. A negative value for c implies that more dense countries in the past now have lower degrees of income inequality. A negative value for c supports the view that these heterogeneous societies have been able to achieve more social mobility over time thereby leading to less inequality. However, the possibility remains that these regions initially had more inequality. A positive value for c implies the opposite, that these societies were not able to break an initial stratification. Instead, a positive coefficient supports the view that past population density created inequities that remained persistent over time.

The unobservable, e, is assumed to be mean zero and have finite but not necessarily equal variance across observations. Consequently, (1) is estimated using heteroskedastic-consistent covariance matrices.

Matrix X contains a set of explanatory variables that control for various historical factors that might influence the distribution of income. These include the European colonial dummies: LISBON, LONDON, MADRID, and PARIS. Each of these takes the

[2] Like DS, the Gini coefficients in this paper are re-scaled to lie between zero and 100.

value one if a country is a former colony of that respective European power and takes the value zero otherwise. Whether or not a country was a colony might have long-run implications for its distribution of income. There might also be differences in income inequality arising from different policies enacted by colonial powers. The control group of countries are those that were never colonies along with South Korea and Taiwan which were Japanese colonies of limited duration. INDEP denotes the natural log of the time since independence to 1995 and is included to determine if the degree of income inequality is associated with how long a country has been in existence. This might be especially relevant for former colonies since colonial income distributions might differ from those under independence and only slowly evolve from the former to the latter.

The following regional dummies are also included: AFRICA (sub-Saharan Africa), WEUR (western Europe), EEUR (eastern Europe), and EASIA (east Asia). AFRICA is included since many African countries also have very skewed distributions of income. EEUR attempts to capture the influence of communism in equalizing incomes. Likewise, western Europe either due to its extensive social programs or because of a differing historical experience from other world regions also has little income inequality. EASIA attempts to account for the fast growth experiences of Japan and the Asian tigers and its possible influence upon income distributions.[3] A Latin American dummy is not included since it would be highly correlated with MADRID + LISBON. Since population density differs across world regions, these regional dummies can also control for regional characteristics that would otherwise be captured by DEN1500 and potentially distort the findings.

LAT comes from Hall and Jones (1998) and denotes the absolute latitude (divided by 90) of a country. More tropical countries were more likely to develop plantation systems of agriculture with their consequent effects on the income distribution as described in Engerman and Sokoloff (1997). Other variables will be added to this baseline specification to check robustness.

An advantage of this approach is that the above variables are most likely exogenous to contemporaneous income distributions and so reverse causation is not likely to be a factor. Although this does not necessarily imply that these controls are determinants of income inequality, they are not effects and so are at least associated with the true determinants if they are not causal factors themselves. A disadvantage of this baseline specification is the lack of more contemporaneous factors that might also affect the distribution of income. Some will later be included in the robustness checks, including current population density. Including current density will ensure that DEN1500 is not

[3] This set of countries includes: Japan, Hong Kong, Singapore, South Korea, and Taiwan. However, this set does not ideally comprise a regional measure as China would be considered in East Asia whereas Singapore is in Southeast Asia. Replacing Singapore with China increases the magnitude of the coefficient on EASIA but does not greatly affect the coefficient upon DEN1500. Moreover, replacing EASIA with ASIA (equals one if country is in Asia and zero otherwise) also does not greatly affect the coefficient on DEN1500.

capturing current demographics as opposed to historical characteristics.

Descriptive statistics are given in Table 1.

Table 1. Descriptive Statistics

Variable	Mean	Median	Standard Deviation
GINI	42.17	41.30	11.08
	46.50	46.07	9.90
DEN1500	7.58	2.46	10.46
	3.60	1.53	5.77
DENNOW	337.06	117.46	945.92
	338.80	79.18	1123.24
DEN0	2.87	0.71	4.41
	1.37	0.42	2.82
DEN1000	4.05	1.46	4.88
	2.64	0.90	4.45
DEN1700	11.57	3.18	16.60
	5.50	1.76	9.54
DEN1900	35.31	10.38	49.94
	17.62	7.69	29.13
INDEP	4.54	4.40	1.14
	4.09	3.83	0.80
LAT	0.29	0.28	0.19
	0.20	0.17	0.13
INC60	7.55	7.49	0.89
	7.31	7.14	0.82
INC60SQ	57.78	56.07	13.47
	54.10	50.95	12.28
YRSOP	0.38	0.27	0.35
	0.28	0.16	0.32
ECORG	3.46	3.00	1.48
	3.43	3.00	1.31
INSTQUAL	6.02	5.68	2.54
	4.96	4.55	2.11
LISBON	0.01		
	0.02		
LONDON	0.32		
	0.44		
MADRID	0.19		
	0.27		
PARIS	0.11		
	0.16		
AFRICA	0.24		
	0.35		
EASIA	0.04		
	0.05		
EEUR	0.07		
	NA		
WEUR	0.18		
	NA		

Notes: Top row denotes value for full sample. Bottom row denotes value for former colony sample. NA - Not applicable in former colony sample.

4. RESULTS

The correlation between GINI and DEN1500 is -0.52 and implies that population density in the past is negatively associated with current income inequality. Figure 1 shows a scatter diagram of DEN1500 versus GINI, which again shows the negative relation between the two. Although formerly low density regions have both high and low inequality today, greater inequality is slightly more prevalent. Regions with higher densities in the past have lower inequality today. The correlation between GINI and current population density is only -0.11. Regressing GINI against a constant and DEN1500 results in a coefficient of -0.55, significant at the 1% level.

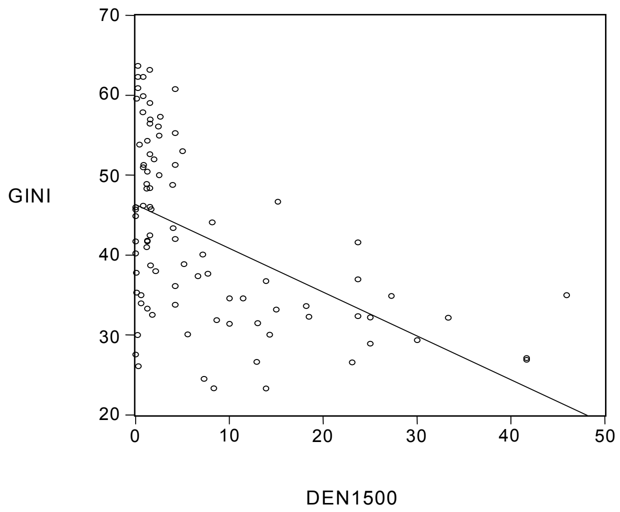

Figure 1. Population Density 1500 AD and Current Income Inequality

The first column of Table 2 presents the results from a baseline specification using the variables described in Section 3 as controls. Coefficients for many of the control variables have the expected signs. Former Spanish and Portuguese colonies have greater inequality than the control group. Western and Eastern European countries have lower inequality. The coefficient on the African dummy is positive and statistically significant. There is no strong evidence that latitude or time since independence is associated with income inequality. The coefficient on DEN1500 falls in magnitude to -0.16 but remains significant at the 5% level. For each additional 10 individuals per square kilometer (roughly one standard deviation), the Gini coefficient is predicted to be 1.6 points lower or one-eighth of a standard deviation. The below results confirm these findings. Population density five centuries ago is negatively associated with income inequality.

Table 2. Full Sample Regressions (GINI is dependent variable)

Column	1	2	3	4	5	6
Constant	38.756***	49.086***	42.891***	-152.367**	-143.581*	-61.731
	(5.251)	(6.564)	(6.622)	(70.223)	(75.042)	(81.598)
DEN1500	-0.162**	-0.169**	-0.142**	-0.178***	-0.207***	-0.167**
	(0.062)	(0.071)	(0.071)	(0.061)	(0.065)	(0.066)
LISBON	17.170***	15.655***	15.667***	15.154***	12.845***	12.952***
	(2.659)	(2.784)	(4.461)	(2.922)	(3.173)	(3.220)
LONDON	2.124	4.194	1.020	3.852	1.874	1.962
	(3.045)	(3.150)	(3.390)	(3.146)	(3.023)	(2.982)
MADRID	6.547*	7.188**	6.045	5.698*	3.147	4.554
	(3.329)	(3.305)	(4.614)	(3.393)	(3.490)	(3.645)
PARIS	-0.510	-1.842	-1.221	0.307	-0.577	0.790
	(5.209)	(4.929)	(5.604)	(4.900)	(4.780)	(4.558)
AFRICA	7.906**	5.948	6.408	8.527**	7.856**	6.385*
	(3.375)	(3.762)	(3.920)	(3.225)	(3.271)	(3.458)
EASIA	-4.186	-3.013	-4.439	-7.002**	-6.961*	-5.971
	(6.169)	(4.007)	(4.272)	(3.486)	(3.717)	(3.934)
EEUR	-13.515***	-10.651***	-13.772***	-14.598***	-12.217***	-8.533**
	(3.624)	(3.812)	(4.460)	(3.477)	(3.975)	(4.182)
WEUR	-7.186**	-4.712	-6.946	-6.119	-5.405	-4.105
	(3.504)	(3.503)	(4.538)	(3.683)	(4.076)	(4.316)
INDEP	0.815	-0.574	0.184	1.357	1.668*	0.827
	(0.899)	(0.958)	(0.847)	(0.921)	(0.980)	(0.846)
LAT	-2.322	-5.871	-7.210	1.334	-1.859	-6.477
	(8.847)	(8.692)	(9.379)	(9.645)	(9.475)	(10.624)
GURRINIT		-0.888				
		(0.535)				
CHRIST			0.017			
			(0.053)			
MUSLIM			0.011			
			(0.055)			
INC60				50.599***	48.410**	28.095
				(18.401)	(19.677)	(21.228)
SQINC60				-3.377***	-3.247**	-1.895
				(1.189)	(1.288)	(1.396)
DENNOW				0.001	0.001	0.0004
				(0.001)	(0.001)	(0.001)
YRSOP					-2.988	-1.688
					(3.888)	(3.755)
ECORG					1.293	1.609*
					(0.927)	(0.890)
INSTQUAL					-0.294	-0.365
					(0.620)	(0.683)
ELF						-4.360
						(3.865)
SECED						-0.060
						(0.059)
R^2	0.558	0.603	0.586	0.606	0.649	0.704
No. of obs.	90	87	84	90	85	78

Notes: Standard errors are in parentheses. *, **, *** indicates significance at 10%, 5%, and 1% levels.

Columns 2 and 3 include other historical characteristics that have potentially influenced the distribution of income but which data is missing for some observations. GURRINIT from Marshall and Jaggers (2000) is the Gurr measure of executive constraints and is averaged from the five years following 1800 or independence, whichever occurs later. GURRINIT attempts to control for institutions at the beginning of a country's history. It is measured on a 1 to 7 integer scale with higher values denoting stronger constraints. A negative coefficient would suggest that countries having institutions that prevented leaders from behaving arbitrarily are more likely to have less inequality today. The coefficient is negative but insignificant. MUSLIM and CHRIST denote the percentage of the population that is Muslim and Christian, respectively. Although these measures are current, they are likely to be indicative of the prevalence of these religions in the past. Neither is strongly associated with income inequality while the coefficient on DEN1500 changes little.

Column 4 begins to add more contemporaneous variables to the baseline specification to better control for current factors that might influence income inequality. However, there are disadvantages to their inclusion since reverse causation is more likely. Moreover, to the extent that influences from DEN1500 affect current income inequality through these other factors, then their inclusion should lower the magnitude of the coefficient on DEN1500. The current variables are: a current measure of population density (DENNOW) as well as the natural log of income per capita in 1960 and its square (INC60 and SQINC60). DENNOW is included to control for current population density and to distinguish its association with income inequality from that between DEN1500 and GINI. Including INC60 and SQINC60 controls for possible nonlinear evolutions of income inequality corresponding to rising average income. They are used to account for the degree of economic development.

The coefficient on DENNOW is nearly zero. There is no evidence that population density is contemporaneously associated with income inequality (at least at the national level). The coefficient on INC60 is positive whereas that on SQINC60 is negative, both findings agree with results from previous studies attempting to ascertain the validity of the Kuznets hypothesis (see Bruno *et al.* (1996) for a survey).[4]

Column 5 adds other contemporaneous variables to control for current policies and institutions. YRSOP denotes the fraction of years between 1950 and 1994 that the economy is considered by Sachs and Warner (1995) to be open. ECORG is from Hall and Jones (1998) although originally from Freedom House. It is an index of integer values from 0 to 5 with higher values denoting greater degrees of capitalism. Both of these variables are included as general measures of economic policy. INSTQUAL is one of the institutional measures from *Political Risk Services* (although taken from Englebert

[4] However, it is not clear that cross sectional data can determine the validity of the Kuznets hypothesis that income inequality first rises and then falls with higher income. This view describes a relation over time within a country whereas the data is taken from one point in time across countries.

(2000)). It is an index from 1 to 10 denoting institutional quality regarding the adjudication of disputes, the enactment of laws, and the replacement of the executive. Higher values denote a greater adherence to a rule of law in these matters.[5] None of these additional variables are significant. The coefficient on DEN1500 increases in magnitude to -0.21.

Column 6 adds a measure of ethnolinguistic fractionalization in 1964 (ELF) with higher values denoting more diversity and human capital (SECED) to the specification. ELF accounts for the degree of ethnic pluralism in society as more pluralistic societies might develop economic cleavages along similar lines. SECED is the enrollment rate in secondary education in 1960. I use SECED since it contains more variation across the sample than does primary enrollment. Ideally, I would like to use a measure of the stock of education in the adult population since this would be less likely than enrollment rates to be affected by current income inequality. I use SECED instead due to greater availability. Neither are significant at conventional levels and the sign on ELF is even negative.

Table 3 conducts robustness checks of some of the above specifications. Columns 1, 2, and 3 use the natural logs of (one plus) the population density variables (denoted by LDEN1500 and LDENNW). Using natural logs places less weight on outliers having extremely high population densities. Although LDEN1500 is not significant in column 1, its sign remains negative and it is significant when additional controls are included as shown in columns 2 and 3. The final three columns only contain observations with income inequality data Deininger and Squire (1996) label as ‘acceptable’. This disproportionately removes poor countries from the sample since they are less likely to have high quality data. The coefficient on DEN1500 remains negative and is significant in two of the three specifications. In addition, these variables are able to account for over 80% of the variance of income inequality across this smaller sample of countries.

As mentioned in Section 2, the relation between population density and income inequality might differ for former colonies where policies and migration were imposed upon an indigenous population. Tables 4 and 5 repeat the above specifications but now include only former colonies or countries such as Nepal and Iran which were never actual colonies but where British authority still held sway. Notwithstanding these exceptions, I refer to this sample as a “former colonial sample.” Thailand, Japan, China, Turkey, and the European countries are removed. EEUR, WEUR, and LONDON are removed from the specification. The control group becomes former British colonies with exceptions such as South Korea and Taiwan (which were Japanese colonies of relatively limited duration). The correlation between GINI and DEN1500 is -0.36 for this sub-sample.

[5] I do not include the Gastil measures of political freedoms or civil liberties since they are less available than is INSTQUAL. However, INSTQUAL does control for factors typically associated with democracies. Replacing INSTQUAL with either the Gastil measure of political freedoms or the Gastil measure of civil liberties or including either with INSTQUAL does not change the conclusions of the paper.

Table 3. Robustness Checks using Full Sample (GINI is dependent variable)

Column	1	2	3	4	5	6
	All data is used			Only acceptable income ineq. data		
Constant	40.113***	-152.255**	-155.032**	40.992***	-178.021**	-142.042*
	(5.658)	(71.055)	(74.178)	(4.303)	(83.896)	(82.920)
DEN1500				-0.077	-0.093*	-0.140**
				(0.049)	(0.051)	(0.061)
LDEN1500	-1.114	-2.564**	-3.543***			
	(0.966)	(1.156)	(1.092)			
LISBON	16.449***	14.047***	9.096**	18.735***	14.959***	13.152***
	(3.337)	(3.606)	(4.156)	(2.852)	(2.779)	(3.099)
LONDON	1.906	3.767	1.780	3.136	2.609	0.244
	(3.197)	(3.003)	(2.853)	(3.021)	(2.980)	(3.030)
MADRID	6.261*	5.273	1.734	10.770***	8.156**	4.595
	(3.562)	(3.600)	(3.886)	(3.460)	(3.255)	(3.511)
PARIS	-0.386	1.891	0.900	11.377**	9.768	4.896
	(5.280)	(4.922)	(4.730)	(5.634)	(6.230)	(8.711)
AFRICA	7.677**	7.864**	6.690*	8.378	7.473	6.486
	(3.434)	(3.228)	(3.459)	(5.047)	(4.786)	(4.632)
EASIA	-5.177*	-8.514**	-8.082**	-3.201	-5.804	-5.269
	(2.993)	(3.346)	(3.857)	(3.117)	(3.643)	(3.864)
EEUR	-13.123***	-13.844***	-11.253***	-8.356**	-11.248***	-8.825*
	(3.623)	(3.403)	(4.039)	(3.975)	(3.836)	(4.789)
WEUR	-7.653**	-5.702	-3.306	-1.695	-3.701	-3.539
	(3.624)	(3.639)	(4.104)	(3.896)	(3.776)	(4.539)
INDEP	0.771	1.504*	2.019**	0.738	0.944	1.430*
	(0.895)	(0.847)	(0.934)	(0.588)	(0.601)	(0.735)
LAT	-3.742	0.493	-3.925	-18.032*	-13.742	-13.636
	(8.697)	(9.029)	(8.897)	(9.256)	(10.880)	(12.294)
INC60		49.971***	51.261***		57.183**	47.549**
		(18.684)	(19.204)		(21.331)	(21.132)
SQINC60		-3.362***	-3.471***		-3.699***	-3.034**
		(1.204)	(1.259)		(1.352)	(1.340)
DENNOW					0.000	0.000
					(0.001)	(0.001)
LDENNW		1.275	1.418			
		(1.161)	(1.130)			
YRSOP			-4.347			-1.923
			(4.039)			(4.371)
ECORG			1.211			1.396
			(0.932)			(1.239)
INSTQUAL			-0.315			-0.922
			(0.631)			(0.568)
R^2	0.551	0.612	0.662	0.810	0.843	0.866
No. of obs.	90	90	85	56	56	53

Notes: Standard errors are in parentheses. *, **, *** indicates significance at 10%, 5%, and 1% levels.

Table 4. Former Colonial Sample Regressions (GINI is dependent variable)

Column	1	2	3	4	5	6
Constant	43.590***	68.777***	59.905***	-197.244**	-245.398**	-80.571
	(11.685)	(12.349)	(13.929)	(95.074)	(111.901)	(133.710)
DEN1500	-0.339	-0.475**	-0.368	-0.533***	-0.593***	-0.540*
	(0.161)	(0.184)	(0.245)	(0.194)	(0.220)	(0.288)
LISBON	15.189***	13.448**	17.214***	8.135*	2.076	7.851
	(4.280)	(5.246)	(5.585)	(4.394)	(4.546)	(6.544)
MADRID	4.748	5.424	7.642	-0.183	-4.065	1.554
	(4.844)	(5.040)	(5.592)	(4.726)	(4.521)	(6.344)
PARIS	-2.599	-6.743	-2.382	-4.252	-4.078	-2.705
	(4.255)	(4.122)	(4.602)	(3.818)	(3.845)	(3.508)
AFRICA	7.507**	3.929	3.937	6.856**	5.241	3.281
	(3.565)	(3.977)	(4.345)	(3.352)	(3.239)	(3.464)
EASIA	-6.570**	-8.628*	-8.355*	-11.799**	-10.514	-9.619
	(3.183)	(4.459)	(4.457)	(5.280)	(6.832)	(6.195)
INDEP	0.118	-4.160	-3.637	1.485	3.927	-0.676
	(3.060)	(2.743)	(3.375)	(2.815)	(2.857)	(3.803)
LAT	1.164	-0.235	-0.609	6.865	0.862	-1.134
	(11.250)	(11.306)	(12.291)	(11.923)	(11.539)	(12.495)
GURRINIT		-1.079				
		(0.717)				
CHRIST			0.014			
			(0.081)			
MUSLIM			-0.003			
			(0.061)			
INC60				65.526***	77.269**	37.417
				(24.420)	(29.389)	(34.532)
SQINC60				-4.477***	-5.344***	-2.598
				(1.597)	(1.976)	(2.362)
DENNOW				0.001	0.001	0.0003
				(0.001)	(0.001)	(0.001)
YRSOP					-6.302	-3.832
					(4.140)	(4.626)
ECORG					1.165	1.428
					(1.106)	(1.040)
INSTQUAL					0.101	-0.176
					(0.937)	(1.014)
ELF						-5.454
						(5.064)
SECED						-0.079
						(0.124)
R^2	0.263	0.355	0.301	0.375	0.449	0.485
No. of obs.	63	60	57	63	58	52

Notes: Standard errors are in parentheses. *, **, *** indicates significance at 10%, 5%, and 1% levels.

Using this smaller sample, the coefficient on DEN1500 is often greater in magnitude (more negative) than its counterpart in the previous two tables.[6] As before, greater population density in the past is associated with less inequality today although the coefficient estimates are less precise. A notable exception occurs when only acceptable data are used in which case the coefficient on DEN1500 is nearly zero (not presented but available upon request.). However, the sample size has been greatly restricted in that only 30 observations remain. The coefficient of determination has also been greatly reduced. These same explanatory variables do less well in explaining income inequality within former colonies relative to a broader group of countries.

One possible explanation for the larger coefficients on DEN1500 is that areas with a small indigenous population attracted settlers who then imported labor leading to a dual society and to more inequality today. Given the removal of the European countries, this possibility exists for a larger percentage of the sample. Nevertheless, the higher standard errors show that this coefficient is less precisely estimated and so such conclusions should be tempered.

Table 5 includes a factor not considered above, the degree of European settlement. EUR1900 and EUR1975 denote the fraction of the population of European descent in 1900 and 1975, respectively. Both are from Acemoglu *et al*. (2001). Countries with little settlement might have greater income inequality since it is more likely that the elite would exploit indigenous populations compared to populations heavily comprised of Europeans. EUR1900 is included to determine if historical settlement patterns are associated with current inequality. EUR1975 is included since Europeans might be more attracted to migrate to countries with less inequality. It is also included to better interpret the coefficient on EUR1900. Since EUR1975 and EUR1900 are correlated, it would not be clear how historical settlement patterns are associated with current income inequality without controlling for the ethnic descent of the current population.[7] The results in Table 5 show little association between historical degrees of settlement (apart from present day ethnicity) and current income inequality. However, countries that now have a greater contemporary degree of European descent have less income inequality. The findings regarding DEN1500 change little, however. Therefore, the above findings regarding former colonies are not entirely due to how local populations might have influenced subsequent European settlement and the consequent effects.

[6] To mitigate the potential for the high income, former British colonies to be driving the findings of this colonial sample, I created the dummy variable CANZUS that equals one for Canada, Australia, New Zealand, and the United States but equals zero otherwise. The coefficient on DEN1500 remained significant and did not greatly decrease in magnitude. Hence, these four countries do not appear to be driving the results. Moreover, the only former Portuguese colony in the sample is Brazil and so the inclusion of LISBON makes it less likely that Brazil is driving the results.

[7] Without EUR1975, the coefficient on EUR1900 is negative and significant.

Table 5. Robustness Checks Using Former Colonial Sample
(GINI is dependent variable)

Column	1	2	3	4	5	6
Constant	33.343*	-188.430*	-203.170	42.612***	-194.465*	-266.277**
	(18.417)	(105.543)	(132.599)	(12.273)	(98.565)	(106.183)
DEN1500	-0.556**	-0.513**	-0.662**			
	(0.209)	(0.250)	(0.265)			
LDEN1500				-1.319	-3.617**	-5.778***
				(1.392)	(1.873)	(1.936)
LISBON	18.919**	13.785*	10.665	15.280***	7.225	-2.952
	(7.608)	(9.969)	(10.106)	(4.462)	(4.795)	(4.719)
MADRID	3.830	-0.068	-1.419	5.261	-0.165	-6.244
	(7.419)	(7.461)	(7.975)	(5.857)	(4.640)	(4.382)
PARIS	-5.217	-6.048	-4.936	-2.127	-2.168	-1.806
	(4.266)	(3.868)	(3.980)	(4.461)	(3.934)	(4.076)
AFRICA	6.427	7.429**	5.026*	8.134**	7.307*	4.316
	(3.439)	(3.266)	(2.961)	(3.728)	(3.753)	(3.851)
EASIA	-3.730	-1.504	-17.637	-6.506**	-11.834***	-10.198
	(3.587)	(38.103)	(39.139)	(3.071)	(4.201)	(6.208)
INDEP	2.876	3.197	4.280	0.392	2.181	5.375
	(4.995)	(4.973)	(5.257)	(3.106)	(2.807)	(2.531)
LAT	22.005*	18.152	12.603	-0.763	1.934	-6.441
	(12.256)	(11.755)	(11.163)	(10.977)	(10.635)	(10.466)
INC60		59.141**	62.220*		63.274**	82.297***
		(27.387)	(34.242)		(25.515)	(28.028)
SQINC60		-3.877**	-4.128*		-4.315**	-5.731***
		(1.794)	(2.351)		(1.666)	(1.880)
EUR1900	18.530	24.065	29.547			
	(21.857)	(21.114)	(20.635)			
EUR1975	-37.478**	-38.557**	-44.048**			
	(17.067)	(16.433)	(18.492)			
DENNOW		-0.001	0.001			
		(0.005)	(0.005)			
LDENNW					1.033	1.335
					(1.307)	(1.108)
YRSOP			-8.151*			-8.027*
			(4.145)			(4.458)
ECORG			1.422			1.100
			(1.163)			(1.125)
INSTQUAL			0.275			0.142
			(1.104)			(0.917)
R^2	0.339	0.419	0.488	0.243	0.360	0.467
No. of obs.	55	55	53	63	63	58

Notes: Standard errors are in parentheses. *, **, *** indicates significance at 10%, 5%, and 1% levels.

Finally, two more general robustness checks are provided. I replaced GINI with the income shares of the 20th, 40th, and 60th percentiles, respectively. Higher values in these quintiles denote that lower income groups are receiving larger shares of income. Unfortunately, these other measures are not available for 25 countries. The results regarding DEN1500 are presented in Table 6 for selected specifications and samples. Generally, results using the full sample remain robust to using any of these income distribution measures. The signs of the coefficients are positive and significant, suggesting that formerly more dense regions now have higher fractions of income going to low income groups. However, results are not always robust for the former colonial sample although the sample size has been cut by over a third from that used in Tables 4 and 5. (Forty observations remain.) Nevertheless, the coefficients on DEN1500 remain positive in all specifications.

Table 6. Robustness Checks using Different Income Distribution Measures: SHARExx denotes income share of xxth percentile

Table, Column	SHARE20	SHARE40	SHARE60
2, 1	0.0005**	0.0011**	0.0012*
	(0.0002)	(0.0005)	(0.0006)
2,5	0.0006***	0.0012**	0.0017**
	(0.0002)	(0.0005)	(0.0007)
3,3[1)]	0.0064**	0.0129*	0.0185*
	(0.0028)	(0.0065)	(0.0104)
4,1	0.0009*	0.0021**	0.0018
	(0.0005)	(0.0010)	(0.0016)
4,5	0.0007	0.0024*	0.0034
	(0.0006)	(0.0013)	(0.0024)
5,1	0.0013**	0.0038***	0.0048**
	(0.0006)	(0.0013)	(0.0022)
5,3	0.0011*	0.0035**	0.0051*
	(0.0006)	(0.0014)	(0.0025)
5,6[1)]	0.0044	0.0103	0.0194
	(0.0054)	(0.0142)	(0.0239)

Notes: Standard errors in parentheses. *, **, *** indicates significance at 10%, 5%, and 1% levels. [1)] denotes natural log of population density variable is used.

Finally, Table 7 considers other periods than 1500. As before, I only present results regarding population density to conserve space. The four periods are the years 0, 1000, 1700, and 1900. The coefficient for population density is generally significant across periods and across specifications. For no specification is the coefficient positive. The coefficients decline in magnitude but that is because population density is generally rising, especially after 1700.

To get a sense of the changing magnitude of the association between past population density across different years and current income inequality holding other factors fixed, one can take the absolute value of the product of the standard deviation of population density for each of the five years with the respective coefficient on population density from the regression. In the baseline specification (Table 2, column 1), the absolute value of this product increases from 1.2 for year 0 to 1.8 for 1700 but then falls to 1.75 for 1900. For Table 2, column 5, this product is lowest for year 1900 at 1.85. The other values are over two. For the analogous specifications in the former colonial sample (Table 4, columns 1 and 5), the patterns are similar. For column 1, the 1900 value is not the lowest but it is lower than the year 1700 product. For the column 5 specification, all these products are over three except for year 1900 which is only 2.1. In all these specifications, therefore, the magnitudes of these products are not necessarily monotonically decreasing but the 1900 products are lower than the 1700 products and sometimes lower than all the rest. This finding supports the view of a weakening association between past population density and current income inequality as more recent historical periods are considered.

Table 7. Robustness Checks with Population Density from Different Periods

Specification	Coefficient for Population Density Variable			
Table, Column	DEN0	DEN1000	DEN1700	DEN1900
2, 1	-0.277	-0.301*	-0.110***	-0.035***
	(0.186)	(0.159)	(0.039)	(0.010)
2,5	-0.479**	-0.524***	-0.141***	-0.037***
	(0.239)	(0.170)	(0.042)	(0.012)
3,3[1)]	-3.971***	-3.891***	-3.685***	-4.565***
	(1.393)	(1.272)	(1.016)	(1.168)
3,6	-0.154	-0.312*	-0.097**	-0.032***
	(0.151)	(0.178)	(0.039)	(0.011)
4,1	-0.710**	-0.386*	-0.236***	-0.061*
	(0.314)	(0.220)	(0.098)	(0.031)
4,5	-1.287***	-0.783***	-0.380***	-0.072**
	(0.309)	(0.287)	(0.124)	(0.030)
5,1	-1.756**	-0.713**	-0.344**	-0.043
	(0.658)	(0.283)	(0.128)	(0.035)
5,3	-2.297***	-0.863**	-0.408**	-0.093
	(0.813)	(0.357)	(0.161)	(0.069)
5,6[1)]	-6.543***	-4.766**	-4.586**	-1.713
	(2.021)	(2.090)	(1.894)	(1.084)

Notes: Standard errors in parentheses. *, **, *** indicates significance at 10%, 5%, and 1% levels. [1)] denotes natural log of population density variable is used.

5. CONCLUSION

This paper examined the association between population density and income inequality. There is no strong association between current population density and income inequality. However, the association between population density in 1500 and contemporary income inequality is negative. More densely populated regions 500 years ago now have less income inequality than do once more sparse areas. This finding is robust to considering other years. This paper does not formally consider reasons for this association, but one possibility is that more dense regions were a sign of more diverse and mobile societies with a nontrivial percentage of the population engaged in activities other than agriculture. A complementary explanation is that such areas enjoyed high agricultural productivity. Not only might this benefit poorer farmers relative to the elite but it could have released workers for these nonagricultural activities thereby adding to social mobility. Over time, this could have resulted in less inequality as more opportunities arose and one group was less able to monopolize power.

However, this association becomes less clear when one considers only former colonies. Although the magnitude of the coefficient increases, its estimate is less precise, implying that the degree of diversity across observations increases. Sparse areas might have led to settlement, small scale farming, and more egalitarian societies whereas dense regions were exploited for labor or tribute. With this added possibility, the relation between historical population density and current income inequality becomes more cloudy, especially since low dense regions might have led to the importation of a slave population thereby increasing income inequality in the long run. Considering colonial policy and slavery and their long-run influences on income distributions are left for future work.

These and other possibilities are certainly worthy of future consideration as we try to identify the determinants of income inequality across world regions. Given that population demographics 500 years ago help explain income inequality today, it behooves researchers to identify other historical determinants and to better understand why these links might be so persistent. I hope this paper is a step in this pursuit.

APPENDIX

Capital letters denotes countries with ‘acceptable’ income inequality data as labeled by Deininger and Squire (1996).

Former Colonial Sample: Benin, Chad, Egypt, GABON, Ivory Coast, Kenya, Lesotho, Madagascar, Malawi, MAURITIUS, Morocco, Niger, Nigeria, Senegal, Seychelles, SIERRA LEONE, SOUTH AFRICA, SUDAN, Swaziland, Tanzania, Togo, Tunisia, Uganda, ZAMBIA, Zimbabwe, BAHAMAS, CANADA, COSTA RICA, DOMINICAN REPUBLIC, EL SALVADOR, GUATEMALA, HONDURAS, JAMAICA, MEXICO, PANAMA, TRINIDAD AND TOBAGO, UNITED STATES, Argentina, Bolivia, BRAZIL, CHILE, COLUMBIA, Ecuador, Peru, Surinam, Uruguay, VENEZUELA, BANGLADESH, India, Indonesia, Iran, Israel, SOUTH KOREA, MALAYSIA, NEPAL, Pakistan, PHILIPPINES, SINGAPORE, SRI LANKA, TAIWAN, AUSTRALIA, Fiji, NEW ZEALAND.

Full Sample includes above countries plus: CHINA, JAPAN, THAILAND, Austria, BELGIUM, BULGARIA, CZECHOSLAVAKIA, DENMARK, FINLAND, FRANCE, WEST GERMANY, Greece, HUNGARY, IRELAND, ITALY, LUXEMBOURG, NETHERLANDS, NORWAY, POLAND, PORTUGAL, ROMANIA, Spain, SWEDEN, Switzerland, TURKEY, UNITED KINGDOM, YUGOSLAVIA

REFERENCES

Acemoglu, D., S. Johnson, and J. Robinson (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, forthcoming.

Alesina, A., and R. Perotti (1996), "Income Distribution, Political Instability, and Investment," *European Economic Review*, 81, 1170-1189.

Alesina, A., and D. Rodrik (1994), "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109, 465-490.

Barro, R. (1999), "Inequality, Growth, and Investment," NBER Working Paper, No. 7738, Cambridge, MA: National Bureau of Economic Research.

Benabou, R. (1994), "Education, Income Distribution, and Growth: The Local Connection," NBER Working Paper, No. 4798, Cambridge, MA: National Bureau of Economic Research.

_____ (1996), "Inequality and Growth," in *NBER Macroeconomics Annual*, eds. by B.S. Bernanke and J.J. Rotemberg, Cambridge and London: MIT Press, 11-74.

Boserup, E. (1965), *The Condition of Agricultural Growth: The Economies of Agrarian Change under Population Pressure*, Chicago: University of Chicago Press.

Bourguignon, F., and C. Morrisson (1990), "Income Distribution, Development, and Foreign Trade: A Cross-Sectional Analysis," *European Economic Review*, 34, 1113-32.

Bruno, M., M. Ravallion, and L. Squire (1996), "Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues," Policy Research Working Paper, No. 1563, The World Bank.

Chakrabarti, A. (2000), "Does Trade Cause Inequality," *Journal of Economic Development*, 25, 1-21.

Chan, S. (1989), "Income Inequality among LDCs: A Comparative Analysis of Alternative Perspectives," *International Studies Quarterly*, 33, 45-65.

Clarke, G. (1995), "More Evidence on Income Distribution and Growth," *Journal of Development Economics*, 47, 403-27.

Crenshaw, E., and A. Ansari (1993), "Dimensions of Social Inequality in the Third

World," *Population Research and Policy Review*, 12, 297-313.
Deininger, K., and L. Squire (1996), "New Ways of Looking at Old Issues: Inequality and Growth," Working Paper, The World Bank, Washington D.C.
Dehghan, F. (1995), "Income Distribution, Overurbanization and Economic Development: Evidence from a Cross-National Study," *Journal of Economic Development*, 20, 15-36.
Diamond, J. (1999), *Guns, Germs, and Steel*, W.W. Norton & Co: New York.
Durlauf, S. (1996), "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1, 75-93.
Engerman, S., and K. Sokoloff (1997), "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies," in *How Latin America Fell Behind*, ed. by S.H. Haber, Stanford University Press: Stanford, CA.
Englebert, P. (2000), *State Legitimacy and Development in Africa*, Lynne Rienner Publishers: Boulder, CO.
Galor, O., and J. Zeira (1993), "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60, 35-62.
Hall, R., and C. Jones (1998), "Why Do Some Countries Produce So Much More Output Per Worker Than Others," Working Paper, Stanford University.
Li, H., L. Squire, and H. Zou (1998), "Explaining International and Intertemporal Variations in Income Inequality," *Economic Journal*, 108, 26-43.
Marshall, M., and K. Jaggers (2000), "Polity IV Project: Political Regime Characteristics and Transitions, 1800-1999," Center for International Development and Conflict Management, University of Maryland, College Park, MD.
McEvedy, C., and R. Jones (1978), *Atlas of World Population History*, Penguin Books: New York.
Midlarsky, M. (1982), "Scarcity and Inequality: Prologue to the Onset of Mass Revolution," *Journal of Conflict Resolution*, 26, 3-38.
Midlarsky, M., and K. Roberts (1985), "Class, State, and Revolution in Central America: Nicaragua and El Salvador Compared," *Journal of Conflict Resolution*, 29, 163-93.
North, D., and R. Thomas (1973), *The Rise of the Western World: A New Economic History*, Cambridge University Press: Cambridge.
Nugent, J., and J. Robinson (2000), "Are Endowments Fate," Working Paper, Department of Political Science, UC-Berkeley.
Perotti, R. (1993), "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies*, 60, 755-76.
Persson, T., and G. Tabellini (1994), "Is Inequality Harmful for Growth: Theory and Evidence," *American Economic Review*, 84, 600-22.
Sylwester, K. (2002), "Democracy and Changes in Income Inequality," *International Journal of Business and Economics*, 1, 167-178.
Weede, E., and H. Tiefenbach (1981), "Some Recent Explanations of Income Inequality," *The International Studies Quarterly*, 25, 255-82.

Mailing Address: Department of Economics, MC 4515, SIU, Carbondale, IL 62901. Tel: 618-453-5347, Fax: 618-453-2717. E-mail: ksylwest@siu.edu

Manuscript received December, 2002; final revision received August, 2003.