The U.S. as a Coastal Nation

Jordan Rappaport and Jeffrey Sachs
Federal Reserve Bank of Kansas City
Center for International Development at Harvard University

This Version: July 2001[†]

U.S. economic activity is overwhelmingly concentrated at its ocean and Great Lakes coasts and at navigable rivers. Economic theory suggests four possible explanations: a present-day productivity effect, a present-day quality-of-life effect, delayed adjustment following a historical productivity or quality-of-life effect, and an agglomeration effect following a historical productivity or quality-of-life effect. Controlling for correlated natural attributes such as the weather and including proximity measures which *a priori* should absorb any quality-of-life effect, linear regressions suggest that the high coastal concentration of economic activity is primarily due to a productivity effect. Extensively controlling for historical economic density suggests that such a productivity effect continues to be operative today.

[†] Thank you for advice and suggestions to Patty Beeson, Jerry Carlino, David Cutler, David DeJong, Claudia Goldin, Edward Glaeser, and seminar participants at Emory University, the University of Pittsburgh, the Federal Reserve Bank of Philadelphia, the 2000 Federal Reserve System Regional Conference, and the 2001 Missouri Economics Conference. A number of individuals have provided excellent research assistance including Nathaniel Baum Snow, Scott Benolkin, Anne Berry, Krista Jacobs, Jason Martinek, and Andrea Zanter. The views expressed herein are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Kansas City or the Federal Reserve System. Comments warmly appreciated. Corresponding email: jordan.m.rappaport@kc.frb.org. Revised versions will be available from www.kc.frb.org/econres/staff/jmr.htm

1 Introduction: Geography Matters

An abundance of rich fertile land and an open frontier uniquely characterize U.S. economic development. Less widely recognized is the extent to which the United States is and has always been a primarily coastal country. Consider Map 1: the shaded area represents the 745 counties with centers within 50 kilometers of an ocean or Great Lakes coast, or within 25 kilometers of a river on which there was commercial transport in 1968. Collectively, these counties account for just 15 percent of the continental U.S. land area but 54 percent of 2000 population and 60 percent of 1998 civilian income. Put differently income per square kilometer of these "coastal" counties is more than eight times that of the remaining "inland" counties.

That the United States with its abundant land remains a primarily coastal nation underscores a basic economic fact: geography matters. In the search to understand the underlying determinants of growth and prosperity, economists have examined a myriad of country attributes ranging from the self evident (e.g. education) to the controversial (e.g. culture). But the role of geography, for the most part, has been neglected.

From a theoretical perspective, modern growth models focus on the accumulation of physical, human, and technological capital which individually or together complement raw labor as the main factors of production. Land, when included, tends to serve as the intensive factor in a traditional sector away from which labor in shifting (Lewis, 1954; Jorgenson, 1961; Ranis and Fei, 1961; Harris and Todaro, 1970; Dixit, 1973; Drazen and Eckstein, 1988). More recently theory has begun to grapple with the issue of space: increasing returns to scale in production, whether direct or via spillovers in technology and human capital, imply a spatial concentration of industry location (Henderson, 1988; Krugman, 1991). While both approaches yield insights, neither addresses the constraints physical geography may place upon economic growth.

This was not always so. Adam Smith in Book 1 of the Wealth of Nations observes the key importance of access to navigable water as an input to the development process:

> As by means of water carriage a more extensive market is opened to every sort of industry than what land carriage alone can afford it, so it is upon the sea-coast, and along the banks of navigable rivers that industry of every kind begins to sub-divide and improve itself, and it is frequently not till a long time after that

> those improvements extend themselves to the inland part of the country.

Thus Smith laments the difficult preconditions for economic growth facing inland Africa and large parts of Russia, Siberia, and Central Asia.

> All the inland parts of Africa, and all that part of Asia which lies any considerable way north of the Black and Caspian Seas, the ancient Scythia, the modern Tartary and Siberia, seem in all ages of the world to have been in the same [economically undeveloped] state in which we find them at present. . . There are in Africa none of those great inlets . . . to carry maritime trade into the interior parts of that great continent.

Recent empirical work using cross-country data confirms that Smith's observation on the role of access to navigable water still holds in the late twentieth century. For example, per capita income has risen considerably faster in coastal regions than in landlocked regions (Gallup and Sachs, 1998); and the share of a nation's population residing within 100 kilometers of an ocean coast is positively correlated with growth in manufactured exports (Radelet and Sachs, 1998). In addition to being landlocked, Central Africa faces several other geographic challenges including weather highly conducive to vectors for parasitic disease transmission and location at a lattitude with low photosynthetic potential contributing to low agricultural productivity (Bloom and Sachs, 1998).

Smith's observation also implicitly underscores the incredibly favorable economic geography enjoyed by the nations of Western Europe. Extensive ocean shorelines and numerous rivers penetrating deep into the interior provide excellent access to navigable water; a temperate climate raises agricultural productivity and helps prevent vector borne disease transmission.

While the United States also enjoys long ocean shorelines and an extensive inland river network, its continental scale nevertheless implies that most of the U.S. land mass lies considerably far from navigable water. Rather than from coastal proximity, an argument can be made that the United States' prosperity derives from its natural resource abundance; indeed its land-based wealth is the stuff of American mythology. Consistent with such an argument, Wright (1990) shows that during the period when the United States moved into a position of world industrial preeminence, the factor content of its net exports was growing increasingly intensive in natural resources. Hence the United States stands out as a possible exception to the importance of access to navigable water in fostering growth.

But such U.S. exceptionalism is misleading. In fact, the United States' economic activity is overwhelmingly concentrated near oceans, Great Lakes, and navigable rivers. Moreover, this concentration has been increasing since the late 19th century.

In what follows we argue that the high concentration of U.S. economic activity at its coasts reflects the productivity effect of access to navigable water. Alternative hypotheses which we explore include that the high concentration derives from a quality-of-life effect from coastal proximity and that it derives from history dependence due to past productivity and quality-of-life effects which no longer hold.

The paper proceeds as follows: Section 2 discusses the theoretical basis for using economic density as a measure of underlying productivity and quality of life and reviews related empirical literature. Section 3 presents some simple empirics illustrating the continual increase in the coastal concentration of economic activity since the late 19th-century and then discusses some demographic differences between coastal and inland counties. Section 4 lays out our econometric specification. Section 5 presents our results on the partial correlates of population density across U.S. counties in 2000. We find a nonlinear relationship between economic density and distance from the coast: density falls off sharply moving slightly inland from ocean and Great Lakes coasts; moving further inland, density continues to fall off but at a more gradual rate. Measuring coastal proximity in ways that a priori should separately identify contributions to productivity and to quality of life shows that the high coastal density follows largely from productivity. Controlling for past economic activity shows that the high coastal density, at least in part, represents current rather than past contributions to productivity. A last section concludes.

2 Theory and Background

For comparing economic outcomes across countries, per capita income serves as a natural measure of welfare; but for comparing economic outcomes across local areas among which individuals and firms can easily move, it does not. A high level of per capita income may reflect high underlying productivity; but it may also reflect "compensation" for undesirable quality-of-life such as unpleasant weather or pollution. Hence it is not clear whether high per capita income represents “good” or “bad” underlying fundamentals.

Rather than per capita income, we use population and employment density as more natural metrics for capturing underlying variations in local productivity and quality of life (Haurin, 1980; Glaeser et. al., 1992, 1995; Ciccone and Hall, 1996). Let a “locality” be defined as a geographic area where people both live and work. Assuming that individuals inelastically supply labor and that there is full employment, population density and employment density will be essentially equivalent. As will be shown below, using either population density or employment density as the dependent variable in our regressions effects nearly identical results.

Consider a locality with a set of attributes that increase the productivity of resident firms. In addition to access to navigable water, some productivity-enhancing attributes might include abundant natural resources, temperate weather, and rule of law. Firms' high productivity increases the marginal revenue product of both labor and capital in turn inducing an inflow of each; moreover, the complementarity between labor and capital implies these inflows are mutually reinforcing (Figure 1a). In a long run steady state, high productivity implies high population density, $\frac{\mathrm{d}L}{\mathrm{d\ productivity}} > 0$.

Similarly, consider a locality with a set of attributes which directly increase the quality of life of local residents. Some quality-of-life enhancing attributes might include ocean vistas, pleasant weather, and low crime. The high quality of life induces an inflow of labor which in turn induces an inflow of capital (Figure 1b). In a long run steady state, $\frac{\mathrm{d}L}{\mathrm{d\ quality\ of\ life}} > 0$. Formal proofs of these are deferred to Appendix A.[1] But the intuition should be straightforward.

Consistent with the idea that people vote with their feet (Tiebout, 1956), population density reveals individuals' preferences over local areas by aggregating the indirect contribution to utility via productivity-driven higher wages with the direct contribution to utility via high quality of life. Map 2, which shows the relative population density of U.S. counties in 2000, represents the result of such a “vote”: the higher the population density, the greater the productivity and quality of life benefits from underlying local attributes. Put differently, Map 2 implicitly defines a population density function isomorphic with a representative agent's indirect utility function over local attributes.

To identify local attributes causing high population and employment density is there-

---

[1] Alternative proofs can be found in Haurin (1980).

Figure 1a: High Productivity Increases Population and Employment

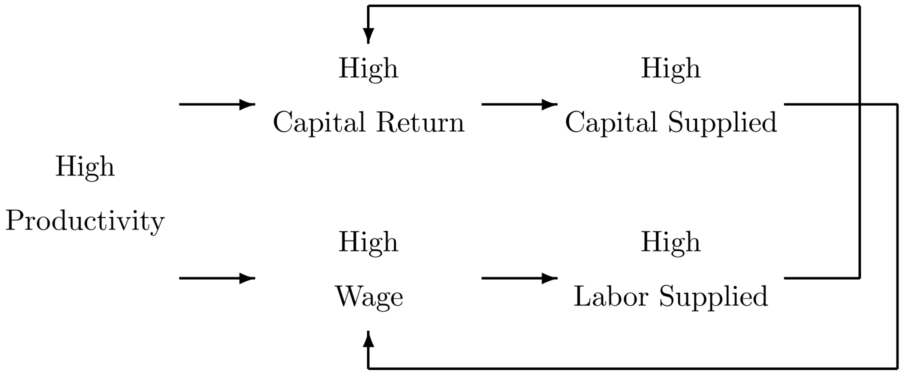
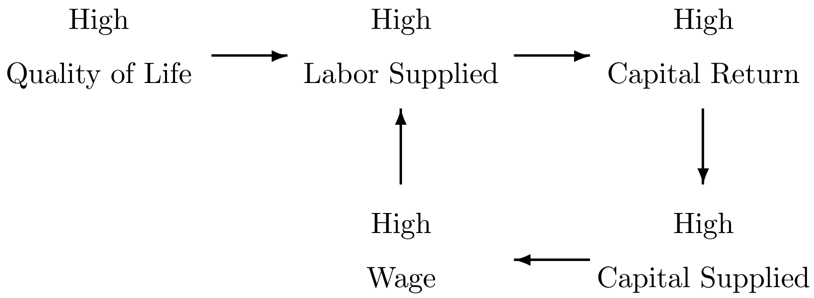


Figure 1b: High Quality of Life Increases Population and Employment



fore to identify attributes positively contributing to local productivity and quality of life. However, few empirical economic studies actually have focused on the partial correlates of population and employment density *levels*. Instead, endogeneity concerns have forced most researchers to focus instead on the partial correlates of population and employment *growth rates*. An important exception is Beeson et al. (1999), who run regressions of county population in 1990 on a vector of natural attributes including several coastal indicator variables. All else equal, they find location bordering the Atlantic or Gulf of Mexico to be associated with a 147 percent increase in population; the presence of a navigable river to be associated with a 14 percent increase in population; and the confluence of two navigable rivers or of a navigable river and the Atlantic or Gulf to be associated with a 130 percent increase in population. Moreover, the first and third of these partial correlations are found to be substantially larger in 1990 than in 1840.

Inherent in population density's aggregating over productivity and quality of life is that it cannot distinguish between the two. In a certain sense, the distinction does not matter. For a representative agent seeking to choose the ideal local attribute mix, population density already "correctly" weights productivity versus quality of life. But from the perspective of economic development, a positive income elasticity of demand implies that individuals living in the United States will tend to place a relatively high value on quality of life. For the coastal concentration of U.S. economic activity to hold lessons for poorer nations, it needs to be shown that such concentration reflects — at least in part — high underlying productivity.

Of course, there is little doubt that living near a coast also may increase quality of life. In addition to overwhelming anecdotal evidence, the quality-of-life contribution from coastal proximity is persuasively established by the compensating differential empirical literature, which values it by the sum of the lower wages individuals are willing to accept and the higher housing prices they are willing to pay to live in coastal areas (Rosen, 1979; Roback, 1982). Controlling for worker-specific and house-specific characteristics, Blomquist et al. (1988) estimate that location adjacent to an ocean or Great Lakes coast lowers annual incomes by \$155 and raises the annual price of housing by \$716; Gyourko and Tracy (1991) estimate these at \$874 and \$1201 respectively.[2] Hence the contribution from coastal location to an average working individual's quality of life would be valued in the range of \$871 (Bloomquist et al.) to \$2075 per year (Gyourko and Tracy). A variation on the compensating differential approach by Stover and Leven (1992) similarly values coastal quality of life at \$721 per year.

Unfortunately, the compensating differential approach proves less useful in valuing coastal proximity's contribution to productivity. In theory, such a contribution to productivity could be valued by the sum of the higher wages, land, and other nontradable input prices firms are willing to pay to operate near a coast. In practice, researchers have not had access to a database on plant-level nontradable inputs with sufficient detail and geographic scope to estimate productivity contributions.[3] Equally problematic is that anonymity re-

---

[2] All values are in 1999 dollars. For Bloomquist et al., the respective standard errors are \$315 and \$54; for Gyourko and Tracy, \$861 and \$661.

[3] For valuing quality of life contributions, housing prices proxy for broader nontradable goods prices. The greater the share of non-housing nontradables in individual's utility functions, the more the compensating differential framework will undervalue contributions to quality of life. For valuing both quality of life and productivity contributions, the compensating differential framework assumes that the law of one price holds for tradable goods (i.e., wages and other nontradable prices are assumed to measure real purchasing power in terms of some tradable numeraire good). However, variations in tradable good prices due to transportation costs are *not* problematic as they should be reflected in compensating variations in nontradable prices.

quirements limit the geographic identification of microdata to local areas with high populations. The Blomquist et al. and Gyourko and Tracy results are based on respective cross-sections of 253 U.S. urban counties and 130 U.S. metropolitan areas. Given the presumption that sparsely settled local areas represent non-random economic outcomes, such a sample selection is likely to strongly bias estimated valuations of geographic attributes.

In Section 3 below, we focus primarily on the combined productivity and quality-of-life effects of coastial proximity as measured by population density. But where the data permit us to do so, we include income and demographic data which together suggest that a large portion of the positive correlation between coastal proximity and population density is due to productivity. In the regression analysis of Section 5, we use different coastal proximity measures to separate productivity and quality-of-life effects.

A second limitation of using population density to measure economic outcomes is the difficulty in distinguishing between present-day contributions to productivity and quality of life versus historical contributions to these. For example, coastal proximity may have greatly increased productivity and quality of life through the end of the 19th-century after which it affected neither. To the extent that adjustment towards the resulting new steady-state spatial distribution were slow, high population density near coasts circa 1900 would remain for a long time thereafter.

Alternatively, historical contributions to productivity and quality of life could effect high current population density via economies of scale. As Fujita and Mori (1996) argue, "[port cities] should have disappeared a long time ago when the original advantage (of cheap water access) became unimportant. Clearly, their continued prosperity can be explained only when we consider the 'lock-in effect' of some self-reinforcing agglomeration forces." So in this case, the hypothesis is that past high population density near coasts subsequently contributed to high and growing productivity levels.

A number of studies have found empirical support for such agglomerative forces. For example, looking across U.S. cities, Glaseser, Kallal, Scheinkman and Shleifer (1992) find employment growth to be positively correlated with measures of local competition and urban variety. Henderson, Kuncoro, and Turner (1995) additionally find employment growth

---

For example, an obvious way by which coastal proximity may increase firms' productivity is by lowering transport costs on its tradable inputs. The compensating differential approach values such savings by the additional amount firms are willing to pay for their nontradable inputs.

to be positively correlated with measures of urban specialization. Looking across U.S. states, Ciccone and Hall (1996) find labor productivity to be positively correlated with average employment density. Looking across U.S. counties, Beeson and DeJong (2000) find population growth to be positively correlated with initial population during the 1940s, 1950s, 1960s, and 1980s.[4]

We, too, find evidence of agglomeration economies. U.S. counties with low-to-medium population density in 1890 grew less than proportionally and U.S. counties with high population density in 1890 grew more than proportionally over the subsequent 110 years. We also find evidence that congestion forces eventually dominate agglomerative ones as suggested by Henderson (1974). U.S. counties with the very highest density in 1890 grew less than proportionally over the subsequent 110 years.

But delayed adjustment and steady-state path dependence cannot explain the coastal concentration of U.S. economic activity. As documented below, such coastal concentration has been increasing since the late-19th century. Moreover, the positive partial correlations between current population density and coastal proximity continue to hold even after extensively controlling for historical conditions.

3 The Coastal Concentration of U.S. Population, Historical and Present

Given the maritime nature of the European colonization of North America, the high coastal concentration of economic activity is hardly new. The establishment of settlements at locations affording easy access to maritime transport allowed for the communication and trade on which the Atlantic economy flourished. In the mid and deep South, tobacco and cotton plantations spread from the ocean coast along the navigable rivers down which they could transport their goods. Indeed the importance of access to navigable water is underscored by the tremendous investments by states and private corporations in the construction of canals during the 1820s and 1830s which in turn facilitated the westward spread of trade and industry (Tanner, 1840; Poor, 1860; Fogel, 1964). That such waterways contributed to productivity is documented by Sokoloff (1988) who shows that U.S. inventive

---

[4] For excellent surveys on the theory and empirics of urban agglomeration economies, see Henderson (1988), Glaeser (1998), and Quigly (1998).

activity over the period 1790 to 1846 as measured by patents per capita is strongly positively correlated with proximity to navigable waterways.

Figure 2 picks up the story in 1870. At that time, the collective population density of all counties with centers within 50 kilometers of an ocean coast is 1.9 times the population density of the contemporary continental United States (i.e., total population of such counties divided by total land for such counties relative to total population divided by total land for all counties excluding those in territories which had yet to be admitted as states). The collective population density of all counties with centers within 50 kilometers of the Great Lakes coast and within 25 kilometers of the river identified by Fogel (1964) as navigable in 1890 are 1.7 and 1.4 times that of the contemporary continental United States.

Figure 2 Panel A shows the subsequent rapid increase in relative population density for the Great Lakes and ocean coastal counties. The Great Lakes coast relative population density exceeds that of the ocean coasts starting in 1890, reaching a level 4.5 times that of the contemporary United States in 1930 before leveling off. From 1930 to 1970, the Great Lakes coast relative population density remains relatively constant; thereafter it rapidly declines, falling to 3.4 in 2000.

The ocean coasts relative population density grows steadily from 1870; it again exceeds the Great Lakes coast relative density starting in 1960, reaching a level 5.0 times that of the contemporary United States in 1970 before leveling off.

Navigable river relative population density, in contrast, shows much less variation over this period; it gradually rises to 1.6 in 1910 and then gradually falls to 1.2 in 2000.

One explanation for the rising concentration of population near coasts is that this is picking up the effects of correlated attributes such as temperate weather, the admission of land-rich interior U.S. states, and the shift in employment out of agriculture and into manufacturing and services. As a first pass at addressing such potential explanations, Figure 2 Panel B shows the partial correlation coefficients of population density regressed on categorical dummies for counties with centers within 50 kilometers of an ocean or Great Lakes coast, or a river navigable 1890. The regressions include 16 weather variables (linear and quadratic mean temperature and precipitation in each of January, April, July, and October), a topography variable (average land gradient, i.e. hilliness), and state fixed effects.

In the 1870 regression, only the navigable rivers dummy admits a positive, statistically significant (at the 0.05 level) coefficient. This gradually increases in magnitude, attaining a high of 0.24 in 1970, which implies that all else equal, such navigable river counties had average 1970 population density 1.3 times that of interior counties. The continual rise in the coefficient on the navigable rivers dummy contrasts with the steady decline in navigable rivers relative population density subsequent to 1910. This dichotomy implies that the falloff in relative population density near navigable rivers was largely due to weather, topography, and state-specific reasons rather than a declining contribution from navigable rivers to productivity and quality of life.

For the Great Lakes, a positive coefficient on the coast dummy first statistically differs from zero in 1890. It steadily rises attaining a high of 0.68 in 1970, which implies that all else equal, counties with centers within 50 kilometers of a Great Lakes coast had population density 2.0 times that of interior counties. Note that here again there is some divergence between the time series of the Great Lakes coast dummy and the corresponding relative population density trend. The latter levels off starting in 1930 and declines rapidly starting in 1970; but the partial correlation coefficient continues to increase through 1970 and thereafter declines more gradually. As with counties proximate to navigable rivers, the contrast suggests that weather and state-specific concerns partially account for the falloff in relative population density near the Great Lakes.

The ocean coast dummy first admits a positive, statistically significant coefficient (at the 0.05 level) in 1910 (the 1890 and 1920 coefficients are significant at the 0.10 level). Starting in 1920, this coefficient rapidly increases, It first exceeds the coefficient on the Great Lakes coast dummy in 1980 and attains a high of 0.72 in 1990, which implies that all else equal, counties with centers within 50 kilometers of an ocean coast had population density 2.0 times that of interior counties.

Figure 3 breaks out the concentration of population density near oceans for each of four different coastal segments: the North Atlantic coast (Maryland north to Maine), the South Atlantic coast (Virginia south to Florida), the Gulf of Mexico Coast, and the Pacific Coast.[5] Figure 3 Panel A emphasizes the extremely high population density of counties with centers within 50 kilometers of the North Atlantic coast. The population density of

---

[5] Justifying a three-way split (Atlantic,Gulf, Pacific), for instance, are different trade opportunities offered by the varying locations. The North Atlantic versus South Atlantic split, on the other hand, follows directly from empirically observed differences between the two.

these counties relative to that of the continental United States is 8.0 in 1870 rising to 14.8 in 1930 thereafter leveling off and eventually falling to 11.2 in 2000.

For counties with centers within 50 kilometers of the South Atlantic, Gulf of Mexico, or Pacific coasts, population density in 1870 is actually below the continental U.S. level. Relative 1870 population density ranges from 0.3 for the Pacific counties up to 0.9 for the South Atlantic counties. The Pacific coastal counties begin to grow rapidly in 1900 and accelerating around 1940. Their relative population density first exceeds 1 in 1910 rising to 2.6 in 1940 and to 5.5 in 1990 before falling slightly to 5.4 in 2000. The South Atlantic and Gulf of Mexico coastal counties' population density first exceed the continental U.S. average in 1890 and 1930, respectively. Their rapid growth begins around 1940. By 2000, the South Atlantic and Gulf of Mexico coastal counties' relative population density have risen to 3.2 and 2.6, respectively.

Figure 3 Panel B shows the partial correlation coefficients of population density with categorical dummy variables for each of the four ocean coasts after controling for weather, topography, and state fixed effects as in Figure 2 Panel B above. For 1870, none of the coefficients is significant at the 0.05 level, though the coefficient on the North Atlantic coast dummy is positive significant at the 0.10 level. The magnitude of the latter coefficient implies that all else equal, the North Atlantic counties have average 1870 population density 1.5 times that of interior counties, far below their actual 1870 relative population density of 8.0. Evidently, weather, topography, and state-specific factors go a long the way towards accounting for the North Atlantic coastal counties' high population density. The coefficient on the North Atlantic dummy steadily rises, first becoming statistically significant at the 0.05 level in 1890 and attaining a high of 1.2 in 1940 before gradually declining to 1.0 in 2000. The implied relative population densities (3.3 in 1940 declining to 2.7 in 2000) remain well below the actual North Atlantic relative population density.

A positive coefficient on the Pacific coast dummy first first statistically differs from zero at the 0.05 level in 1910. It grows rapidly over the period 1910 to 1930 and again from 1940 to 1970 attaining a high of 1.1 in 1970 before gradually declining to 1.0 in 2000. As with the North Atlantic coastal counties, the Pacific Ocean coast accounts for only part of the high population density of proximate counties. Starting in 1940 and especially since 1970, weather, topography, and state-specific factors account for a large portion of the Pacific coastal counties' growth.

A positive coefficient on the Gulf of Mexico coast dummy first statistically differs from 0 at the 0.05 level in 1920. Thereafter, it grows fairly steadily attaining 0.9 in 2000, which implies a relative population density of 2.5, nearly identical to the Gulf of Mexico coast counties' actual relative population density.

The coefficient on the South Atlantic coast dummy actually remains negative through 1950, statistically differing from zero at the 0.05 level in 1900 and 1920. For 1960 through 2000, the dummy admits a positive but not statistically significant coefficient. Its relatively small magnitude implies that all else equal, counties with centers within 50 kilometers of the South Atlantic Coast should have relative population density no more than 1.2 times that of interior counties, far below their actual relative population density over this period. So at the South Atlantic coast, weather, topography, and state-specific factors almost entirely account for high observed population density.

The rising relative population density at ocean and Great Lakes coasts documented in Figures 2 and 3 contradicts a delayed adjustment explanation for the present day coastal concentration of economic activity. However, it remains possible that the rising trend represents the agglomeration of initially dense economic activity or the structural shift out of agriculture. To show that such explanations cannot account for the increasing coastal concentration of economic activity will rely on the multivariate analysis of Section 5.

Table 1 recaps the very high density of recent economic activity shown in Map 1 and Figures 2 and 3. Three alternative measures of economic density — population, employment, and labor income — all effect the same ranking of the coastal categories: the North Atlantic counties, with economic density from twenty to thirty times that of inland counties, followed by (in decreasing order) the Pacific, the Great Lakes, the South Atlantic, the Gulf of Mexico, and the navigable rivers. A fourth measure of economic density, capital income, effects nearly the same ranking except that capital density in the South Atlantic coastal counties exceeds that in the Great Lakes coastal counties.[6] (Note that in contrast to the 1890-based navigable river definition used in Figures 2 and 3 above, navigable rivers hereafter are defined as rivers on which there was actual navigation in 1968 based on Southern Illinois University at Carbondale, 1968). Even in the least dense coastal category, navigable

---

[6] Capital income is the sum of dividends, interest, and rent received by individuals based on where they live. On the one hand, capital income intensity may be a poor measure of local productivity as the source of the capital income may be located elsewhere. On the other hand, the productivity with which individuals can invest non-locally may vary across localities, e.g., due to better information.

rivers, economic density ranges from three to four times that of inland counties.

Higher per capita income in the coastal counties is suggestive that the coastal concentration provides from a productivity advantage rather than a quality-of-life one. For all the coastal counties, 1998 per worker annual labor income averaged \$36,747 versus \$28,702 for the remaining inland counties. By separate coast, 1998 per worker income averaged from \$43,331 for the North Atlantic coastal counties down to \$31,436 for the South Atlantic coastal counties. Compensating differential theory suggests that controlling for worker quality, wages measured in terms of tradable goods should be positively correlated with productivity but negatively correlated with quality of life.[7] However, the higher worker incomes may reflect higher worker skills near coasts rather than higher worker productivity. In particular, 22.2% of adults in the coastal counties had at least a Bachelor's degree versus just 17.7% of adults in inland counties (both figures are for 1990).[8]

A second demographic characteristic that suggests that the coastal concentration of economic activity may derive from a productivity advantage is the high percentage of the coastal population that is working age. In choosing where to locate, households presumably place a higher weight on productivity relative to quality of life the higher their desired labor force participation. Desired household labor force participation, in turn, should be proportional to the percentage of individuals who are working age. In 1990, 62.6% of the coastal population was between the ages of 18 and 64 in 1990 versus 60.9% of the inland population. For the most part, the high percent of the coastal population that is working age reflects a low percent of the coastal population that is age 0 to 17: as low as 23.1% for the North Atlantic coastal counties verus 26.6% in the inland counties. Inland counties would seem to have a comparative advantage in child rearing (e.g., in terms of prices and opportunity costs). For the Pacific coastal counties, the high percent of the population that is working age additionally reflects a low percent of the population 65 or older: 10.5% verus 12.6% for the nation as a whole.

---

[7] As the theoretical result is for wages measured in terms of a tradable good numeraire, it is *not* necessary to adjust for higher non tradable goods prices. Indeed, the equalization of individual utility across localities implies that firms' productivity should not affect real wages measured in terms of the price of the typical consumption bundle.

[8] One possible explanation is that coastal proximity disproportionately increases the productivity of highly educated individuals. Alternatively, the higher incomes of highly educated individuals may cause them to disproportionately value a quality-of-life attribute associated with coastal proximity.

While the high per worker incomes and high percent of the population that is working age are suggestive, neither establishes that high productivity underlies the coastal concentration of U.S. economic activity. Similarly, the increasing coastal concentration of economic activity is suggestive but does not establish that it is underpinned by a current rather than a historical functional relationship. Instead, we turn to econometric analysis to identify the various ways in which coastal proximity might affect productivity and quality of life, past and present.

4 Econometric Specification

We assume a data generating process for steady-state economic density in locality i,

$$\begin{array}{ccccccccccc} L^{\star}_{i,t} & = & \boldsymbol{\beta}_t\,(& \mathbf{x}_i &) & + & \mu_t & + & \nu_i & + & \xi_{i,t} \\ \Uparrow & & & \Uparrow & & & \Uparrow & & \Uparrow & & \Uparrow \\ \begin{array}{c}\text{Log}\\ \text{Steady-}\\ \text{State}\\ \text{Density}\end{array} & & & \begin{array}{c}\text{time}\\ \text{invariant}\\ \text{included}\\ \text{attrib.}\end{array} & & & \begin{array}{c}\text{time}\\ \text{intrcpt}\\ \\ \\ \end{array} & & \begin{array}{c}\text{time}\\ \text{invariant}\\ \text{excluded}\\ \text{attrib.}\end{array} & & \begin{array}{c}\text{idiosync.}\\ \text{attrib.}\\ \\ \\ \end{array} \end{array} \tag{1a}$$

The vector $\mathbf{x}_i$ includes measures of locality i's coastal proximity along with measures of correlated geography attributes such as weather and topography. The exogenous nature of these attributes eliminates a reverse-causal interpretation of partial correlations.

As the effects of coastal proximity may change with tastes and technology, steady-state economic density is assumed to be a time-varying function, $\boldsymbol{\beta}_t(\cdot)$, of the time-invariant $\mathbf{x}_i$. Steady-state economic density is modeled as additionally depending on a time intercept term, μ_t; non-modeled time-invariant attributes, ν_i; and time-varrying idiosyncratic attributes, $\xi_{i,t}$.

In practice, steady-state economic density is not observable. To proxy for it, we use current economic density. Hence we estimate,

$$\begin{aligned} L_{i,t} &= \mathbf{x}_i'\boldsymbol{\beta}_t + \mu_t + \underbrace{\nu_i + \xi_{i,t} + \left(L_{i,t} - L^{\star}_{i,t}\right)}_{\text{error term}} \\ &= \mathbf{x}_i'\boldsymbol{\beta}_t + \mu_t + \epsilon_{i,t} \end{aligned} \tag{1b}$$

The difference between steady-state and current economic density gets subsumed in the error term. This difference is likely to be non-trivial. Observed persistent U.S. local population and employment flows suggest that economic density converges only very slowly towards its steady-state level (Rappaport 2000). With slow convergence and the time-varying dependence of steady-state economic density on the exogenous attributes, $\mathbf{x}_i$, the latter will in general be correlated with $\left(L_{i,t} - L^\star_{i,t}\right)$. Therefore, $\widehat{\boldsymbol{\beta}}_t$ estimated using (1b) will be a biased. Intuitively, $\widehat{\boldsymbol{\beta}}_t$ captures a combination of the past and current dependence of economic density on $\mathbf{x}_i$.

So a possible interpretation of non-zero estimated coefficients on the $\mathbf{x}_i$ in (1b) is that instead of measuring a current functional relationship, $\partial L^\star_{i,t}/\partial \mathbf{x}_i$, they capture a past functional dependence that no longer holds in the present. For instance, a particular attribute, x^k_i, may have had a positive impact on steady-state economic density at some time in the past, $\beta^k_{t-1} > 0$, but not in the present, $\beta^k_t = 0$. If so, the growth regression,

$$\mathrm{d}L_{i,t} = \mathbf{x}_i'\mathrm{d}\boldsymbol{\beta}_t + \mathrm{d}\mu_t + \mathrm{d}\epsilon_{i,t} \tag{2}$$

should yield $\widehat{\mathrm{d}\beta}^k_t < 0$.

More difficult is distinguishing between (1a) and an alternative data generating process characterized by steady-state path dependence,

$$\begin{array}{ccccccccccccc}
L^\star_{i,t} & = & \boldsymbol{\Gamma}_t\,(& L_{i,t-1} & , & \mathbf{x}_i &) & + & \mu_t & + & \nu_i & + & \xi_{i,t} \\
\Uparrow & & & \Uparrow & & \Uparrow & & & \Uparrow & & \Uparrow & & \Uparrow \\
\text{Log} & & & \text{Lagged} & & \text{time} & & & \text{time} & & \text{time} & & \text{idiosync.} \\
\text{Steady-} & & & \text{Density} & & \text{invar.} & & & \text{intrcpt} & & \text{invar.} & & \text{attrib.} \\
\text{State} & & & & & \text{incld.} & & & & & \text{excld.} & & \\
\text{Density} & & & & & \text{attrib.} & & & & & \text{attrib.} & &
\end{array} \tag{3a}$$

In practice, this is usually estimated by,

$$\begin{aligned}
L_{i,t} &= \gamma_t L_{i,t-1} + \mathbf{x}_i'\boldsymbol{\delta}_t + \mu_t + \underbrace{\nu_i + \xi_{i,t} + \left(L_{i,t} - L^\star_{i,t}\right)}_{\text{error term}} \qquad (3\text{b}) \\
&= \gamma_t L_{i,t-1} + \mathbf{x}_i'\boldsymbol{\delta}_t + \mu_t + \epsilon_{i,t}
\end{aligned}$$

The specification (3b) is structurally equivalent to the growth regression, (2), with the addition of initial density as a right-hand-side variable. Put differently, (2) constrains the coefficient, γ, on lagged density in (3b) to be one.

The interpretation of the estimated $\widehat{\gamma}$ is problematic. Suppose that (1a) is the true data generating process. The slow adjustment of economic density towards its steady state implies that running (3b) should estimate $\widehat{\gamma} > 0$ despite that $\partial L^\star_{i,t}/\partial L_{i,t-1} = 0$. Reinforcing this upward bias is that $L_{i,t-1}$ contains information on the time invariant excluded attributes proxied by ν_i. Indeed, the greater the variance of ν_i, the greater the tendency for (3b) to estimate $\widehat{\gamma} \approx 1$.[9] Hence finding that $0 < \widehat{\gamma} \leq 1$ does not imply that steady-state economic density depends on lagged economic density.

On the other hand, under (1a) there is no reason to expect that (3b) should estimate $\widehat{\gamma} > 1$. Such a finding can be taken as sufficient evidence of history dependence, $\partial L^{\star}_{i,t}/\partial L_{i,t-1} > 0$. Subtracting lagged density, $L_{i,t-1}$, from both sides of (3b) gives a conditional divergence interpretation of $\widehat{\gamma} > 1$: all else equal, places with higher economic density grow at a quicker rate.[10] Conversely, $\widehat{\gamma} < 1$ is often interpreted as evidence of conditional convergence: all else equal, places with higher economic density grow at a slower rate.

Even more problematic is interpreting the $\widehat{\boldsymbol{\delta}}_t$ estimated from (3b). With conditional convergence, these are often interpreted as measuring the sign of the structural relationship, $\partial L^{\star}_{i,t}/\partial x_i$. Justifying such an interpretation is an identifying assumption that density grows at a rate linearly proportional to its gap from its steady state: $\mathrm{d}L_{i,t} = -(1-\gamma)\cdot\left(L_{i,t} - L^{\star}_{i,t}\right)$. With $\gamma < 1$ and steady-state density generated by a linear version of (1a), the $\widehat{\boldsymbol{\delta}}_t$ should equal $(1-\gamma)\cdot\widehat{\boldsymbol{\beta}}_t$ and so have the same sign as $\widehat{\boldsymbol{\beta}}_t$.

In practice, our estimates of $\boldsymbol{\delta}_t$ are almost always identical in sign and similar in magnitude to our estimates of $\mathrm{d}\boldsymbol{\beta}_t \equiv \boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}$ from the growth specification (2). This is true even after allowing for a richer, nonlinear dependence of current on lagged density. Hence we interpret the $\widehat{\boldsymbol{\delta}}_t$ as capturing the *change* in effect of $\mathbf{x}_i$ on steady-state economic density. Such an interpretation corresponds exactly to the partial derivative of (3a) with respect to $\mathbf{x}_i$.

Note that assuming (1a) is the true data generating process with $\mathrm{d}\beta^k_t$ small, (3b) will tend to estimate $\widehat{\delta}^k_t \approx 0$. Such a finding in no way implies that $\beta^k_t = 0$. On the other hand, regardless of whether (1a) or (3a) is the true data generating process, identically signed estimates $\widehat{\beta}^k_t \gtrless 0$ from (1b) and $\widehat{\delta}^k_t \gtrless 0$ from (3b) can be taken as sufficient evidence that $\partial L^{\star}_{i,t}/\partial x^k_i$ is also of this same sign.

Because any omitted variables with a distribution affected by geography induce spatial correlations among the error terms, we use a generalization of the Huber White heteroskedastic consistent estimator based on Conley (1999) to report standard errors robust to such a spatial structure. For county pairs the Euclidean distance between which is beyond

---

[9] This latter source of bias is emphasized by Islam (1995) and Caselli, Esquivel, Lefort (1997).

[10] Because of slow convergence and excluded attributes, $\widehat{\gamma} > 1$ does not imply the presence of agglomeration economies, $\partial L^{\star}_{i,t}/\partial L_{i,t-1} > 1$.

a certain cutoff, $\overline{d}$, we impose that the covariance between error terms is zero. Within this distance, we impose a (weakly) declining weighting function, $g(\text{distance})$, on the covariance between errors. In essence, this amounts to allowing for a spatially-based random effect. Dropping time subscripts,

$$\begin{aligned} E\left(\varepsilon_i\right) &= 0 \\ E\left(\varepsilon_i \varepsilon_j\right) &= g\left(\text{distance}_{ij}\right)\rho_{ij} \\ \hat{\rho}_{ij} &= e_i e_j \end{aligned} \tag{4}$$

$$\begin{aligned} g\left(\text{distance}_{ij}\right) = 1 &\quad \text{for distance}_{ij} = 0 \\ g\left(\text{distance}_{ij}\right) = 0 &\quad \text{for distance}_{ij} > \overline{d} \\ g'\left(\text{distance}_{ij}\right) \leq 0 &\quad \text{for distance}_{ij} \leq \overline{d} \end{aligned} \tag{5}$$

Herein, we assume the weighting on the covariance between error terms falls off quadratically as the distance between county centers increases with 200 kilometers as the cutoff beyond which we impose zero covariance. So $g(\cdot) = 1 - \left(\frac{\text{distance}_{ij}}{200}\right)^2$.

Note that the error specification in (4) and (5) reduces to the Huber White heteroskedastic consistent estimator for standard errors when $\overline{d}$ equals zero; it reduces to a group-based random effect estimator for standard errors with a non-Euclidean distance measure and a one-zero step specification for $g(\cdot)$.[11]

5 Empirical Results: The Coastal Determinants of Economic Density

To reject the hypothesis that coastal proximity does not affect economic density, Maps 1 and 2 and Figures 2 and 3 presented above should be sufficient. Our purpose in pursuing multivariate regression analysis, instead, is to better describe the magnitude and nature of coastal proximity's effect on economic density. In particular, we are interested in distinguishing between contributions to productivity versus contributions to quality of life; and in distinguishing between whether such contributions only occurred historically versus whether they continue up through the present.

---

[11] Equivalent to the "cluster" option in Stata.

5.1 Base Specification

For our base specification, we measure coastal proximity using six different variables. Three separate coastal dummy variables are set equal to 1 for counties with centers within 50 kilometers of an ocean coast, of a Great Lakes coast, or of a river on which there was commercial navigation in 1968. The first two of these dummies are identical to those used in Figures 2 and 3; the navigable river definition for the third dummy differs in that it is based on usage in 1968 rather than 1890 (Map 5 versus Map 4). Our interest in identifying current contributions to productivity along with the very low minimum standards defining "navigability" in 1890 motivate our choice to define navigability based on more recent usage. However, the result is a selection bias in that potentially navigable rivers around which there is insufficient economic density to support commercial navigation will be excluded from our sample. Such criticism is anyway made moot by the fragility of the partial correlations with respect to navigable river proximity.

The remaining 3 coastal proximity variables measure "far distance" to each of the nearest ocean coast, the nearest Great Lakes coast, and the nearest river on which there was commercial navigation in 1968. Whereas the coefficient on the dummy variables captures a discrete effect of coastal proximity, the coefficient on the far distance measures captures the corresponding gradient with which economic density falls off moving away from the coast from beyond a certain threshold distance inland. More specifically, far distance' to oceans and Great Lakes equals $\log(1+\text{distance from county center to coast})-\log(51)$ for counties with centers more than 50 kilometers distant from the respective coast, zero otherwise; far distance to navigable rivers equals $\log(1+\text{distance from county center to navigable river})-\log(26)$ for counties with centers more than 25 kilometers distant from such a river, zero otherwise.[12]

Finally, our base specification includes the 16 weather variables (linear and quadratic mean temperature and precipitation in each of January, April, July, and October) and topography variable (average land gradient) used in Figures 2 and 3 above. Except as a robustness check, we do not include state fixed effects.

Table 4 Column 1 reports results from regressing $\log(1+\text{population density in 2000})$ on

---

[12] The additive terms are included so that the coefficients on the dummy variables correctly measure the coastal "discrete" effect. As specified, a county with center just above 50 kilometers from an ocean or Great Lakes coast (just above 25 kilometers from a navigable river) will have a fitted density level of one.

our base specification. The coefficient on the ocean coast dummy is positive and significant at the 0.05 level. Its magnitude implies that location at an ocean coast (i.e., with county center no more than 50 kilometers away) is associated with population density 2.3 times that of location elsewhere. The elasticity of population density with respect to far distance from an ocean coast is negative and significant at the 0.05 level. Its magnitude implies that a county with center 200 kilometers from an ocean coast is associated with population density 0.7 that of a county with center 50 kilometers from an ocean coast. The top panel of Figure 4 illustrates.

Our base specification shows the Great Lakes coasts to exert a gradient effect on population density but not a level one. The elasticity of population density with respect to far distance from a Great Lakes coast is negative and significant at the 0.05 level. Its magnitude implies that a county with center 200 kilometers from an ocean coast is associated with population density 0.5 that a county with center 50 kilometers from a Great Lakes coast. But the coefficient on the Great Lakes dummy is not statistically significant. The middle panel of Figure 4 illustrates.

The base specification coefficient on the navigable river dummy is positive and significant at the 0.05 level; the elasticity of population density with respect to far distance from a navigable river is negative and significant at the 0.10 level. Location at a navigable river (i.e., with county center no more than 25 kilometers away) is associated with population density 1.4 times that of location elsewhere. A county with center 100 kilometers from a navigable river is expected to have population density 0.9 that of a county with center 25 kilometers from a navigable river. The bottom panel of Figure 4 illustrates.

Table 4 Columns 2 and 3 report results from regressing log(1+population density in 2000) on the coastal proximity measures using two alternative specifications. Column 2 controls for nothing else; Column 3 includes state fixed effects in addition to the weather and topography controls. For the ocean and Great Lakes coastal proximity measures, the pattern of coefficient signs and statistical significance is identical across the three specifications. For the navigable river proximity measures, either the dummy coefficient is positive significant or the elasticity coefficient is negative significant or both across the three specifications. Compared to the base specification, not controlling for weather and topography moderately lowers the positive coefficient on the ocean coast dummy and sharply raises (in absolute value) the ocean far distance elasticity coefficient.

The coastal proximity measures successfully account for a large part of the variation in population density across counties. On their own, the six proximity measures account for 43 percent of such variation. Additionally controlling for weather and topography, the base specification accounts for 51 percent of such variation. Additionally controlling for state fixed effects results in only a small increase in explanatory power to 56 percent. For comparison, controlling only for weather and topography (but not coastal proximity) accounts for 45 percent of the variation. Controlling only for state fixed effects accounts for 41 percent of the variation.

The base specification partial correlations of population density with coastal proximity are moderately robust. Table 4 Columns 4 through 6 show quite similar results obtain from using $\log(1 + \text{employment density in 1998})$ as the regression dependent variable. (Though explanatory power is from 7 to 8 percentage points lower.) Table 9 Column 1 shows the results from varying the distance delimiting the dummy variable from far distance. Using 25 kilometers rather than 50 as the cutoff, the positive coefficient on the ocean coast dummy no longer statistically differs from 0; but the coefficient on ocean coast far distance increases (in absolute value). Using 75 kilometers as the cutoff, the coefficient on ocean coast far distance statistically differs from 0 only at the 0.10 level; but the coefficients on the ocean coast dummy and Great Lakes far distance both increase (in absolute value). Using 100 kilometers as the cutoff, the coefficient on ocean coast far distance no longer statistically differs from 0. Using 35 or 50 kilometers rather than 25 as the cutoff, the coefficient on navigable river far distance no longer statistically differs from 0.

5.2 Controlling for History

To address the possibility that these partial correlations are picking up a past functional relationship that no longer holds, Table 5 reports results from regressions on the coastal proximity measures of the change in population density and the level of population density extensively controlling for historical population density. Hereafter, all regressions control for weather and topography but not state fixed effects.

Table 5 Column 1 has as its dependent variable, $\log(1 + \text{population density in 2000}) - \log(1 + \text{population density in 1890})$.[13] The regression admits a significant positive coefficient

---

[13] An advantage of using the change in $\log(1 + \text{population density})$ rather than the change in $\log(\text{population density})$ is that the former implicitly underweights counties with low population densities. We

on the ocean coast dummy and a significant negative coefficient on Great Lakes far distance. In other words, the population density "premium" to location on an ocean coast became much larger between 1890 and 2000. And the falloff in population density moving away from the Great Lakes (from at least 50 kilometers inland) became steeper between 1890 and 2000. Figure 5 illustrates. Table 5 Column 4 reports analogous results using the change in population density between 1930 and 2000 as the regression dependent variable. For this period as well, the premium to an ocean coast location grew and the falloff with respect to Great Lakes far distance became steeper. In addition, for this latter period the change in the elasticity of population density with respect to ocean coast far distance positively differs from zero at the 0.05 level. In other words the falloff in population density moving away from an ocean coast became less steep between 1930 and 2000.

As discussed in the specification section above, these increases in the partial correlations between population density and coastal proximity may be spurious in the sense that coastal proximity may have caused high initial population density levels which in turn caused high population growth. Consistent with such a story, Beeson and DeJong (2000) find the U.S. counties' population growth rates during the 20th century are positively correlated with their initial population levels.

We, too, find evidence that the evolution of counties' population density is characterized, in part, by divergence. Table 5 Column 2 shows the results from regressing log(1 + population density in 2000) on a constant and log(1 + population density in 1890) where the latter variable is entered as a seven-segment spline. The top panel of Figure 6 illustrates. The 45 degree (dashed) line represents current population density being linearly proportional to initial population density (i.e., with unitary coefficient). Points below this line represent current population density being less than proportional to initial population density. Points above this line represent current population density being more than proportional to initial population density. So for counties with very low initial population density, population density in 2000 was linearly proportional to population density in 1890.[14] For

---

believe this is desirable given that idiosyncratic shocks (e.g., the migration choices of only a few individuals) could otherwise disproportionately affect results.

[14] The magnitude of the (unreported) constant from the regression reported in Table 5 Column 2 implies that U.S. counties' population density grew at an average annual rate of 0.9 percent over the period 1890 to 2000. For the regression reported in Table 5 Column 5, the constant implies that U.S. counties' population density grew at an average annual rate of 0.4 percent over the period 1930 to 2000.

counties with moderately low initial population density (corresponding the the interval labeled “a”), population density in 2000 was less than proportional to population density in 1890. For counties with moderately high initial population density (corresponding to the interval labeled “b”), population density in 2000 was more than proportional to population density in 1890. Roughly speaking, for counties with initial population density in the combined interval, “a” + “b”, the evolution of population density is characterized by divergence: the higher initial population density counties grew more quickly than the lower initial population density counties.[15]

But we also find evidence that the evolution of counties’ population density is characterized, in part, by convergence. For counties with very high initial population density (corresponding to the interval labeled “c”), population density in 2000 was less than proportional to the population density in 1890. This latter result is consistent with Carlino and Chatterjee (2001) and Chatterjee and Carlino (2001), who document convergence of employment density across *metropolitan area* counties over the period 1950 to 1996.

Together the coefficients on the initial population density spline suggest that for most counties, agglomerative forces dominate congestion forces. But for very high population density counties, congestion forces dominate agglomerative forces. The regression reported in Table 5 Column 5 and the associated illustration in the bottom panel of Figure 6 show that similar dynamics characterize the evolution of population density over the period 1930 to 2000. Controlling for weather, topography, and coastal proximity as in Table 5 Columns 3 and 6, the coefficients on the initial population density spline imply nearly identical relationships to those shown in Figure 6.

For present purposes, the main point is that agglomerative forces cannot account for the increase in coastal population density. Table 5 Column 3 shows the results from regressing

---

[15] More strictly speaking, growth is increasing in initial population density if the associated coefficient is greater than one. So for instance, the lower initial population density portion of interval “a” (the 20-to-50 percentile initial population density counties) could be said to be characterized locally by convergence and the upper initial population density portion of interval “a” (the 50-to-80 and 80-to-90 percentile initial population density counties) could be said to be characterized locally by divergence. However, we argue that it is more informative to use language characterizing whether the relationship between current and initial population density lies above or below the 45 degree line. This cross-sectional comparison between current and lagged population density is reminiscent of the Markov-transition methodology introduced by Quah (1993a, 1993b, 1996a, 1996b).

population density in 2000 on the coastal proximity variables while controlling for the 1890 historical population density spline. The coefficient on the ocean coast dummy remains positive and significant at the 0.05 level; its magnitude hardly changes. The coefficient on Great Lakes far distance remains negative and significant at the 0.05 level, though its magnitude is somewhat diminished from the change regression in Table 5 Column 1. These same results would seem to rule out a sectoral shift explanation for the increasing coastal concentration of population. To the extent that past agricultural activity may have been inversely correlated with coastal proximity, the historical density spline should control for this.[16]

The findings of a positive, statistically significant coefficient on the ocean coast dummy and a negative, statistically significant coefficient on Great Lakes far distance are extremely robust to alternative specifications of the change regressions and the level regressions controlling for historical population density. Such alternatives include using 1930 rather than 1890 as the initial population density year (Table 5 Columns 4 and 6), varying the distance delimiting the adjacency dummy from far distance (Table 9 Columns 2, 3a, 4, and 5a), and augmenting the historical population density control set also to include measures of adjacent counties' historical population density (Table 9 Columns 3b and 5b).[17] So even after extensively controlling for initial population density, the positive correlation between population density and coastal proximity increased throughout much of the 20th century.

An additional result in Table 5 is that controlling for initial population density, the coefficient on ocean coast far distance positively differs from zero at the 0.05 level. In other words, the decline in population density moving away from the ocean coast from at least 50 kilometers inland became less steep between 1890 and 2000. This decrease in the population density gradient proves to be even larger (in absolute value) when controlling for harbor proximity as discussed in Section 5.4 below. We defer our interpretation until there.

---

[16] To the extent that past agricultural activity may have been orthogonal to coastal proximity, there is no reason to expect that the shift out of agriculture would have led to an increase in coastal concentration. However, such an increase would be expected if coastal proximity were a productive input for the industries into which activity was shifting. This latter possibility is consistent with the interpretations herein.

[17] These "historical adjacency controls" are the lowest population density of an adjacent county, the highest population density of an adjacent county, and the mean population density of all adjacent counties. These are entered as linear and quadratic terms of $\log(1 + \text{adjacent population density})$ for a total of 6 historical adjacency controls.

5.3 Population Density Level and Change by Separate Ocean Coast

Table 6 reports results from regressions analogous to those in Tables 4 and 5 except that coefficients are allowed to vary based on a specific coast to which a county is closest: the North Atlantic, the South Atlantic, the Gulf of Mexico, the Pacific, or the Great Lakes. All regressions include the weather and topography controls, proximity to navigable rivers, and coast-specific intercepts. As the coefficients on the Great Lakes proximity measures are quite similar to those in Tables 4 and 5, we do not report them here.

Table 6 Column 1 shows the results from the straight "level" regression (i.e., without controlling for initial population density). A positive partial correlation between population density and coastal proximity holds at each of the four ocean coasts. However, the nature of this concentration varies. For counties closest to the North Atlantic, the coefficient on the dummy for counties with centers within 50 kilometers of the coast does not statistically differ from zero. Instead, population density falls off with elasticity -1.1 moving away from the North Atlantic coast from at least 50 kilometers inland. Such elasticity implies that a county with center 200 kilometers from the North Atlantic coast would be expected to have population density 0.22 times that of a county with center 50 kilometers from the North Atlantic coast (Figure 7, top panel, solid line). For counties closest to the South Atlantic, the Gulf of Mexico, and the Pacific, a positive coefficient on the respective ocean coast dummy positively differs from zero at the 0.05 level. The coefficient magnitudes imply that counties with a center within 50 kilometers of the respective coast on average have population density 1.90 times, 2.40 times, and 4.11 times the population density of counties located further inland.

Table 6 Columns 2 and 4 show corresponding change regressions over the respective periods, 1890 to 2000 and 1930 to 2000. Over both periods, the falloff of population density with distance from the North Atlantic coast became more steep; and the premium to location within 50 kilometers of the South Atlantic and the Gulf of Mexico increased. Figure 7 illustrates. The same results also hold running level regressions of population density in 2000 controlling for historical population density in 1890 and 1930.[18]

---

[18] Historical population density is entered with the same seven-part spline used in Table 5 but allowing the associated coefficients to vary by nearest coast.

5.4 Productivity Versus Quality of Life

As laid out in the theory section above, the positive partial correlations between population density and coastal proximity may derive either from a productivity effect or from a quality-of-life effect. To help disentangled between these two, Tables 7 and 8 report regressions of population density on coastal proximity measures that we believe, a priori, affect productivity or quality of life but not both.

For oceans and Great Lakes, we augment our base specification to include analogous proximity measures to harbors. Our prior is that harbors are likely to raise productivity but not quality of life. We also augment the specification to include the ratio of a county's shoreline to its total area. Our prior is that shoreline measures access to recreational and scenic amenities and so primarily impacts quality of life but not productivity.

An ideal harbor measure would be all coastal geological formations affording shelter for seagoing vessels above a certain size threshold. In practice, we identify harbors as a subset of actual seaports included in the World Port Index (U.S. Naval Oceanographic Office, 1971). This classifies seaports by four size categories — very small, small, medium, and large — based on several applicable factors including area, facilities, and wharf space. We define "harbors" as medium or large seaports. To minimize the possibility of reverse causality, we further exclude from our harbor measure any seaports that rely on constructed breakwaters or tide gates rather than natural barriers for shelter. Map 3 shows the resulting "natural harbors" as well as the excluded medium and large seaports relying on constructed shelter. A selection bias remains in that geological formations affording the necessary shelter but that did not actually develop into seaports will be excluded from our harbor measure. As a robustness check to address this selection bias, we will show that our results are robust to alternative harbor measures encompassing the excluded small and very small naturally-sheltered seaports.

For navigable rivers, we augment our base specification to include distance to the nearest "major" river. Major rivers are defined as a superset of navigable rivers to include the longest North American rivers as well as shorter rivers which connect lakes to the ocean (see data appendix). Map 5 illustrates. Our prior is that controlling for proximity to major rivers, any residual correlation of economic activity with proximity to navigable rivers is likely to be picking up a productivity effect. On the other hand, to the extent that population density is correlated with the presence of major rather than navigable rivers, the underlying mechanism may be either productivity (e.g., hydroelectric power) or quality of life (e.g., fishing).

Table 7 Column 1 shows the results from regressing log(1+population density in 2000) on this expanded coastal proximity specification. For oceans, a positive coefficient on the natural harbor dummy and a negative coefficient on far distance to the nearest natural harbor are both significant at the 0.05 level. The coefficient on the former is especially large relative to the analogous coefficient in the base specification in Table 4 (1.20 verusus 0.69) implying that counties with centers within 50 kilometers of a natural harbor have average population density 3.3 times that of counties further away. Controlling for proximity to natural harbors cuts in half the magnitude of the positive coefficient on the ocean coast dummy; it remains statistically significant at the 0.10 level. And the coefficient on far distance to the nearest ocean coast no longer statistically differs from zero. The positive coefficient on the ocean natural harbor dummy is robust to alternatively defining harbors to include small and very small seaports (Table 10 Column 1).[19]

For the Great Lakes, a negative coefficient on far distance to the nearest natural harbor is negative and statistically significant at the 0.05 level while far distance to the nearest Great Lakes coast no longer statistically differs from zero. This negative gradient with respect to Great Lakes natural harbor far distance is robust to alternatively defining Great Lakes harbors to include small and very small ports (Table 10 Column 1). Note however that these more expansive harbor definitions effect a negative coefficient on the resulting Great Lakes harbor dummy.

We can reject that these partial correlations between population density and proximity to natural harbors just derive from history dependence. In both the 1890-to-2000 change regression and in the 2000 level regression controlling for the population density spline in 1890, the coefficient on the ocean natural harbor dummy remains positive significant (Table 7 Columns 2 and 3). In both of these regressions as well as in the analogous 1930-to-2000 regressions, the coefficient on far distance to the nearest Great Lakes natural harbor remains negative significant (Table 7 Columns 2 to 5). In other words, the productivity contribution from being located adjacent to an ocean natural harbor increased over the period 1890 to 2000. And the falloff in the contribution to productivity moving away from natural harbors

---

[19] However, proxying for harbors as the ocean mouths of either major river or rivers navigable in 1890 results in coefficients on the dummy that do not statistically differ from zero.

(from at least 50 kilometers away) became more steep over the periods 1890 to 2000 and 1930 to 2000. Again such results are robust to alternatively defining harbors to include small and very small seaports (Table 10, Columns 2, 3a, 4, 5a) as well as to controlling for the historical population density of adjacent counties (Table 10, Columns 3b and 5b).

Together these results suggests that a large part of the concentration of economic activity near ocean and Great Lakes coasts stems from a productivity rather than a quality-of-life effect. Moreover, such a productivity effect appears to have increased subsequent to 1890. Reinforcing such a conclusion, in both the level and change regressions the coefficient on ocean and Great Lakes shorelines is either negative or does not statistically differ from zero.

On the other hand, major rivers dominate navigable ones in accounting for high population density. A positive coefficient on the dummy for counties with centers within 25 kilometers of a major river statistically differs from zero at the 0.05 level and a negative elasticity of population density with respect to far distance to major rivers statistically differs from zero at the 0.10 level. The coefficient on the navigable river dummy no longer statistically differs from zero. And the statistical significance of the coefficient on navigable river far distance falls to the 0.10 level. Moreover, the lack of a statistically significant coefficient on this variable in any of the change regressions or level regressions controlling for initial population density in Columns 2 through 5 leaves open the possibility that the level result is due to history dependence. Overall, we cannot reject that the concentration of population density near navigable rivers has nothing to do with their navigability per se.

A final pair of results highlighted by Table 7 concerns the changing relationship between population density and proximity to the ocean coast. Even after controlling for distance to ocean natural harbors, a positive, statistically significant coefficient remains on the ocean coast dummy in the change regressions and in the level regressions controlling for historical population density. So population density is increasing immediately adjacent to ocean coasts, regardless of proximity to harbors. The same regressions also admit a positive, statistically significant coefficient on ocean coast far distance. So moving inland from the increasing population density counties with centers less than 50 kilometers from an ocean coast, population density is falling off at a slower rate than in the past. Both results are extremely robust to alternate natural harbor controls (Table 10) as well as to excluding such natural harbor controls (Table 5).

The increase in population density immediately adjacent to ocean coasts after controlling for harbor proximity suggests that individuals may be increasingly valuing the quality-of-life contributions from such location. The decrease in the population density gradient moving further inland is harder to interpret. One possibility is that individuals are willing to endure longer "commutes" to obtain ocean coast quality-of-life benefits. Such an interpretation is consistent with the higher congestion implied by rising population density immediately adjacent to the ocean coast. An alternative interpretation is that beyond some cutoff distance from the ocean coast, any associated benefits (quality of life or productivity) are smaller than in the past. Consistent with this latter interpretation, the decrease in the gradient is greater the larger the cutoff delimiting the ocean coast dummy from "far distance" (Table 9).

Table 8 reports results analogous to those in Table 7 but allowing coefficients to vary by nearest coast. The coefficients on the Great Lakes proximity measures are quite similar to those in Table 7 and so are not shown. In addition to the topography and weather controls, all regressions also include proximity to navigable rivers and coast-specific intercepts.

The level regression reported in Table 8 Column 1 has a positive coefficient on each of the ocean natural harbor dummies, significant at the 0.05 level for the North Atlantic, Gulf of Mexico, and Pacific coasts and significant at the 0.10 level for the South Atlantic coast. At the North Atlantic coast, a negative elasticity of population density with respect to far distance from the nearest natural harbor is also significant at the 0.05 level. In the 1890-to-2000 change regression and the 2000 level regression controlling for 1890 population density, a positive coefficient on the natural harbor dummy significant at either the 0.05 or 0.10 level occurs at each of the South Atlantic, Gulf of Mexico, and Pacific coasts (Table 8 Columns 2 and 3). So at each of the ocean coasts, there is evidence of a positive productivity effect. Moreover, except at the North Atlantic coast, this productivity effect appears to have increased subsequent to 1890.

The regressions reported in Table 8 also give some support to their being a quality-off-life effect from coastal proximity. In the level regression (Table 8 Column 1), the coefficient on the Gulf of Mexico and Pacific coast dummies is positive significant . Also, a positive coefficient on Pacific Coast shoreline statistically differs from zero at the 0.05 level. There is some evidence that such a quality-of-life effect may be increasing with time. In the change regressions and the level regressions controlling for historical population density

(Table 8 Columns 2 to 5), a positive coefficient on the South Atlantic ocean coast dummy is significant at the 0.10 level or less for the period 1930 to 2000 and a positive coefficient on the Gulf of Mexico ocean coast dummy is significant at the 0.05 level for both the 1890-to-2000 and 1930-to-2000 periods. Additionally, three of these four historical regressions result in a positive significant coefficient on either the Pacific Coast ocean dummy or Pacific Coast shoreline.

6 Conclusions

Economic density in the United States is overwhelmingly concentrated at its ocean, Great Lakes, and navigable river coasts. This concentration has been increasing throughout much of the 20th century. Extensively controlling for historical conditions suggests that the coastal concentration captures a present-day contribution to productivity and quality of life. The stronger partial correlations of density with proximity to harbors rather than with the coast per se suggests that the contribution to productivity may be more important than the contribution to quality of life.

The actual mechanism by which coastal proximity increases productivity remains an open question. Obviously, proximity to harbors lowers transportation costs for many tradable goods. To the extent that interregional and international trade make up a continually rising share of U.S. output, it is unsurprising that access to low transportation costs may be increasingly advantageous.

On the other hand, our intuition is that lower transportation costs are not sufficient to account for the overwhelming magnitude of the coastal concentration of U.S. economic activity. We hypothesize that an informational advantage may accompany the trade effected by access to low-cost transport. Alternatively, low-cost transport may interact with agglomeration or quality-of-life effects. Alternatively, it may just be that we have not done a good enough job controlling for initial conditions. Future research needs to address these and other possibilities.

For developing nations, our results reinforce the present consensus on the importance of openness to trade in promoting economic growth. Countries blessed with large ocean ports should stand to benefit from increasing world trade. For other countries, “getting a port” may not be a policy option. Development policy in these countries, instead, needs to take account of what is likely to be a large productivity disadvantage. Of course, doing so requires a better understanding of how coastal proximity affects productivity. So again, more research is needed.

Appendices

A The Compensating Differential Framework

Assume a large number of localities across which there is high labor and capital mobility. In a long run spatial steady state, no individual should be able to increase their utility by moving to a different locality; nor should any firm be able to increase their profitability by doing so. Any variations in exogenous local attributes which affect utility and profits must be offset by compensating wage and nontradable price differentials.

The equating of utility levels across localities is captured by,

$$V\left(p, w; \text{quality of life}\right) = \left\{\max_{c,n}\ u\left(c, n; \text{quality of life}\right) \text{s.t. } c + pn \leq w\right\} = \overline{V} \tag{A.1}$$

$$\begin{aligned} u_c\left(\cdot\right) &> 0;\ u_{cc}\left(\cdot\right) < 0 \\ u_n\left(\cdot\right) &> 0;\ u_{nn}\left(\cdot\right) < 0 \\ u_{\text{quality}}\left(\cdot\right) &> 0;\ u_{c,\text{quality}}\left(\cdot\right) = u_{n,\text{quality}}\left(\cdot\right) \end{aligned}$$

Here, $V(\cdot)$ represents an indirect utility function with the price of land services, p, and the wage level, w, as its arguments, and quality of life as a shift parameter. The underlying (direct) utility function, $u(\cdot)$, is increasing in consumption of a tradable good, c, and nontradable land services, n. With the tradable good as numeraire and normalizing the per capita quantity of inelastically supplied labor to one, individuals face the budget constraint that their tradable consumption plus their expenditure on land services not exceed the wage rate. The first two sets of derivative restrictions just establish that utility is strictly increasing and concave with respect to both the tradable and nontradable goods. The third set of derivative restrictions establishes that a higher quality of life indeed raises individual utility but that it does not alter the relative utility tradeoff between the tradable and nontradable goods.

The equal profit condition is captured by,

$$\Pi\left(w, \bar{r}; \text{productivity}\right) = \left\{\max_{K,L}\ F\left(K, L; \text{productivity}\right) - wL - \left(1+\bar{r}\right)K\right\} = \overline{\Pi} \tag{A.2}$$

$$F_K\left(\cdot\right) > 0;\ F_L\left(\cdot\right) > 0;\ F_{\text{productivity}}\left(\cdot\right) > 0$$

$\Pi\left(\cdot\right)$ represents a firm profit function which, given local wages and an exogenous interest rate, is the maximized value of firm production less its wage and interest bill. The derivative assumptions establish that the marginal products of capital and labor always remain positive and that higher productivity indeed raises output.

Normalizing the quantity of land to one, and assuming a unit flow of land services from each unit of land, a representative locality's resource constraint gives,

$$nL = 1 \tag{A.3}$$

Note that for the representative locality, L measures both population and population density. Generalizing to localities with different (fixed) quantities of land, L should be interpreted only as population density. For the analysis which follows, the key theoretical results are that

$$\frac{\mathrm{d}w}{\mathrm{d\ productivity}} > 0; \qquad \frac{\mathrm{d}p}{\mathrm{d\ productivity}} > 0; \qquad \frac{\mathrm{d}L}{\mathrm{d\ productivity}} > 0 \tag{A.4}$$

$$\frac{\mathrm{d}w}{\mathrm{d\ quality\ of\ life}} = 0; \qquad \frac{\mathrm{d}p}{\mathrm{d\ quality\ of\ life}} > 0; \qquad \frac{\mathrm{d}L}{\mathrm{d\ quality\ of\ life}} > 0 \tag{A.5}$$

To establish that $\frac{\mathrm{d}w}{\mathrm{d\ productivity}} > 0$, recognize that $\Pi(\cdot)$ is a profit function. Hence its derivatives with respect to input prices are negative: $\Pi_w(\cdot) < 0$ and $\Pi_{\bar{r}}(\cdot) < 0$. Using the envelope theorem, we know that

$$\begin{aligned} \frac{\mathrm{d}\Pi(\cdot)}{\mathrm{d\ productivity}} &= \frac{\partial\Pi(\cdot)}{\partial\ \mathrm{productivity}} + \Pi_w(\cdot)\frac{\mathrm{d}w}{\mathrm{d\ productivity}} = 0 \\ &= F_{\mathrm{productivity}}(\cdot) + \Pi_w(\cdot)\frac{\mathrm{d}w}{\mathrm{d\ productivity}} = 0 \end{aligned} \tag{A.6}$$

By assumption $F_{\mathrm{productivity}}(\cdot) > 0$. Hence $\frac{\mathrm{d}w}{\mathrm{d\ productivity}} > 0$.

To establish that $\frac{\mathrm{d}p}{\mathrm{d\ productivity}} > 0$, recognize that $V(\cdot)$ is an indirect utility function. Hence its derivative with respect to its resource constraint will be positive, $V_w(\cdot) > 0$, and its derivative with respect to the prices of utility arguments will be negative, $V_p(\cdot) < 0$. Taking the total derivative of $V(\cdot)$, setting this equal to zero, and rearranging gives,

$$\frac{\mathrm{d}p}{\mathrm{d\ productivity}} = -\frac{V_w()}{V_p()}\frac{\mathrm{d}w}{\mathrm{d\ productivity}} > 0 \tag{A.7}$$

To establish that $\frac{\mathrm{d}w}{\mathrm{d\ quality\ of\ life}} = 0$, totally differentiate $\Pi(\cdot)$ and rearrange.

To establish $\frac{\mathrm{d}p}{\mathrm{d\ quality\ of\ life}} > 0$, the envelope theorem gives,

$$\begin{aligned} \frac{\mathrm{d}V(\cdot)}{\mathrm{d\ quality}} &= \frac{\partial V(\cdot)}{\partial\ \mathrm{quality}} + V_p(\cdot)\frac{\mathrm{d}p}{\mathrm{d\ quality}} = 0 \\ &= u_{\mathrm{quality}}(\cdot) + V_p(\cdot)\frac{\mathrm{d}p}{\mathrm{d\ quality}} = 0 \end{aligned} \tag{A.8}$$

By assumption $u_{\mathrm{quality}}(\cdot) > 0$. Hence $\frac{\mathrm{d}p}{\mathrm{d\ quality\ of\ life}} > 0$.

Finally, to show that population density rises with increases in productivity and quality of life, $\frac{\mathrm{d}L}{\mathrm{d\ productivity}} > 0$ and $\frac{\mathrm{d}L}{\mathrm{d\ quality\ of\ life}} > 0$. By the economy resource constraint, (A.3), this is equivalent to showing that per capita land consumption drops with such changes, $\frac{\mathrm{d}n}{\mathrm{d\ productivity}} > 0$ and $\frac{\mathrm{d}n}{\mathrm{d\ quality\ of\ life}} > 0$.

$$\frac{\mathrm{d}V(\cdot)}{\mathrm{d\ productivity}} = u_c(\cdot)\frac{\mathrm{d}c}{\mathrm{d\ productivity}} + u_n(\cdot)\frac{\mathrm{d}n}{\mathrm{d\ productivity}} = 0 \tag{A.9}$$

$$\frac{\mathrm{d}V(\cdot)}{\mathrm{d\ quality}} = u_{\text{quality}}(\cdot) + u_c(\cdot)\frac{\mathrm{d}c}{\mathrm{d\ quality}} + u_n(\cdot)\frac{\mathrm{d}n}{\mathrm{d\ quality}} = 0 \tag{A.10}$$

Individual utility maximization gives,

$$p = \frac{u_n(\cdot)}{u_c(\cdot)} \tag{A.11}$$

Suppose that $\frac{\mathrm{d}n}{\mathrm{d\ productivity}} > 0$. By (A.9) it follows that $\frac{\mathrm{d}c}{\mathrm{d\ productivity}} < 0$. $u(\cdot)$ is such that $\frac{u_n(\cdot)}{u_c(\cdot)} = p$ must fall. But this violates that $\frac{\mathrm{d}p}{\mathrm{d\ productivity}} > 0$. Hence $\frac{\mathrm{d}n}{\mathrm{d\ productivity}} < 0$. The same argument using (A.10) establishes that $\frac{\mathrm{d}n}{\mathrm{d\ quality\ of\ life}} < 0$.

An important caveat to the partial derivatives in (A.4) and (A.5) is that several rely on the exclusion of land from the production function, $F(\cdot)$. When land is included in the production function as in Roback (1982) and Gyourko and Tracy (1989, 1991), the derivative of the output-denominated wage with respect to quality of life, $\frac{\mathrm{d}w}{\mathrm{d\ quality\ of\ life}}$, is negative: in order to attain their reservation level of profits, firms pay a lower output-denominated wage as compensation for the higher output-denominated land price. With land excluded from the production function, the derivative with respect to quality of life of the output-denominated wage is zero but the derivative with respect to quality of life of the Hicksian, consumption-denominated real wage is negative. More importantly, the positive derivative of population density with respect to productivity may not follow. Higher productivity causes an outward shift in both firms' and individuals' demand for land services (due, respectively, to an increase in the marginal product of land and the income effect of higher output denominated wages). Together with the resulting increase in the price of land services, higher productivity may cause the actual aggregate quantity of land services purchased by firms and the per capita quantity of land services purchased by individuals to either increase or decrease. When land is absent from the production function, the price effect dominates the income effect so that per capita land service consumption drops and hence population must increase. But if firms increase their aggregate use of land, then even a decrease in per capita land service consumption may not be sufficient to prevent a decrease in population. Hence the aggregate framework used herein may not be appropriate for examining the contribution from attributes which primarily increase the productivity of land-intensive industries.

Data

Our choice of counties as the unit of observation is motivated by the near constancy of their borders across time. To a first approximation, therefore, these borders can be considered historically determined and therefore exogenous relative to most data generating processes which may be studied. Constant borders allow for intertemporal comparisons between geographically fixed areas; municipal and metropolitan area borders, in contrast, show considerable variation across time. To be sure, occasional changes in county borders do occur. Most frequently such changes take the form of the splitting of a county into two or more counties. Wherever possible, we have recombined such "split" counties to allow for intertemporal comparisons. A second type of adjustment we have made is the combining of counties to achieve geographic contiguity. Particularly in Virginia, there exist a number of "independent cities" completely surrounded by counties from which they are formally separate; we have merged such "independent cities" into their surrounding counties. Washington D.C. is included as a county equivalent. Because of their unique geographic locations, we exclude counties within the states of Hawaii and Alaska.

Ocean and Great Lakes coasts and county boundaries are based on the 1:1.25 million ArcUSA Map constructed and distributed by ESRI Corporation (www.esri.com). For each county, the ESRI software package ArcView was used to calculate the distance to the nearest shoreline from the county's centroid (a mathematical approximation of "the center" of an irregular polygon). Note therefore that even counties with long coastal borders will generally have a positive distance to the coast.

Population and land area data are derived from various years of the U.S. Department of Commerce's decennial census. These are disseminated in electronic form from several different sources listed in the bibliography. Employment and income data listed in Table 1 is from the U.S. Department of Commerce's Bureau for Economic Analysis *Regional Economic Information System*, Tables CA-05 and CA-25. Age and Education data listed in Table 1 is from the 1990 Decennial Census, Summary Tape File 3C.

"Navigability" of rivers is based on a 1968 academic study of inland waterway commercial traffic and requires a minimum channel depth of nine feet (Southern Illinois University at Carbondale, 1968). The inclusion of man-made canals within the navigable river category highlights the potential endogeneity of proximity to to navigable rivers coasts. More generally, maintaining the navigability of natural "navigable" rivers is a challenge requiring the continual attention of the U.S. Army Corps of Engineers. To the extent that the funding for the maintenance of navigability may be correlated with population density, a reverse causal link will exist from population density to navigable river proximity. Even so, the prerequisite of a river basin with free flowing water and at least local navigability place an upper bound on the degree of such reverse causality. A second concern is selection bias: any navigable or potentially navigable river on which there was no commercial traffic in 1968 would be excluded from our navigable classification. Because both of these concerns would bias away from zero the coefficients on the navigable river coast proximity variables, these coefficients need to be interpreted with considerably more caution than those on the ocean coast and Great Lakes coast proximity variables.

Major rivers (regardless of navigability) are made up of all rivers in the 1:25 million North

America map from ESRI Corporation combined with a few navigable rivers which were not included. The ESRI map is constructed to include the longest rivers as well as shorter rivers which connect lakes to the ocean. The ESRI map additionally seeks “that the visual density of the rivers reflect, to a degree, the amount of flowing water present in a region.” (ESRI, email correspondence with author, 10/01/98.)

Rivers navigable 1890 are based on the map of commercially navigated waterways in 1890 included in Fogel (1964). This map was then used to edit the “Major Water” shape file (distributed by ESRI, produced by Geographic Data Technology Incorporated) to remove river portions that are not part of the Fogel set. We further removed a handful of rivers included by Fogel that were only locally navigable (i.e. that did not afford the ability to navigate continuously to an ocean or Great Lakes coast).

Natural ports represent a subset of the seaports included in the World Port Index (U.S. Naval Oceanographic Office, 1971). This catalogs all U.S. Great Lakes and ocean seaports as well as some ports on navigable rivers. As the physical prerequisites for establishing a port on a navigable river are minimal, we exclude ports located more than 100 kilometers from an ocean or Great Lakes coast. Proximity to the more inland ports will instead be captured by the navigable river measures. The 100 kilometer boundary allows cities that are usually considered to be seaports to be classified as such (e.g., Houston, 20 km inland; Philadelphia, 43 km inland; Portland OR, 85 km inland) while excluding cities more commonly considered to be river ports (e.g., Albany, 172 km inland; Memphis, 521 km inland). The World Port Index classifies seaports by four size categories — very small, small, medium, and large — based on several applicable factors including area, facilities, and wharf space. We define “harbors” as medium or large seaports. To minimize the possibility of reverse causality we further exclude from our harbor measure any seaports that rely on constructed breakwaters or tide gates for shelter. The resulting “natural harbors” instead are distinguished by being sheltered from the wind and sea by virtue of a location within a natural coastal indentation or in the protective lee of an island, cape, or other natural barrier or by being located on a river adjoining the ocean. For the robustness check in Table 10, we continue to exclude seaports more than 100 kilometers inland or that rely on constructed shelter.

Our weather and altitude variables are borrowed from Mendelsohn, Nordhaus, and Shaw (1994) who derive these based on observations from 5,511 meteorological stations over the period 1951 through 1980. The temperature variables represent the average over these 30 years of mean daily temperature in the months of January, April, July, and October. The precipitation variables represent average monthly precipitation in these same months. The actual county observations are interpolated values for county geographic centers based on data from surrounding weather stations.

Slope percent is derived from the National Resource Inventory (NRI) database published by the U.S. Department of Agriculture. This suveys land characteristics at almost 800,000 sites across the

United States. For a given site, slope percent represents the vertical rise in altitude of the site over the site's horizontal length multiplied by 100. The actual county observations are a land-weighted average of sites within a county using NRI provided weights.

Latitude and longitude values used to compute the spatially-robust standard errors correspond to the location of county centroids as determined by the ArcView software package.

Bibliography

Bloom, David and Jeffrey Sachs (1998). "Geography, Demography, and Economic Growth in Africa." *Brookings Papers on Economic Activity*, 2:1998.

Beeson, Patricia E., David N. DeJong, and Werner Troesken (1999). "Population Growth in U.S. Counties, 1840-1990." Mimeo, University of Pittsburgh, (April).

Beeson, Patricia E. and David N. DeJong (2000). "Divergence." Mimeo, University of Pittsburgh, (November).

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort (1996). "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics." *Journal of Economic Growth* 1, 3 (September), pp. 363–389.

Carlino, Gerald and Satyajit Chatterjee (2001). "Employment Deconcentration: A New Perspective on America's Postwar Urban Evolution." Federal Reserve Bank of Philadelphia Working Paper 01-4, (March).

Chatterjee, Satyajit and Gerald Carlino (2001). "Aggregate Metropolitan Employment Growth and the Deconcentration of Metropolitan Employment." *Journal of Monetary Economics*, Forthcoming.

Ciccone, Antonio and Robert E. Hall (1996). "Productivity and the Density of Economic Activity." *American Economic Review* 86, 1 (March), pp. 54 – 70.

Conley, Timothy G. (1999). "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92, 1 (Sept.), pp. 1-45.

Dixit, Avinash (1973). "Models of Dual Economies." In J. A. Mirrlees and N. H. Stern, eds., *Models of Economic Growth*, New York: Wiley & Sons.

Drazen, Allan and Zvi Eckstein (1988). "On the Organization of Rural Markets and the Process of Economic Development." *American Economic Review* 78, 3 (June), pp. 431–443.

Fogel, Rober William (1964). *Railroads and American Economic Growth: Essays in Econometric History.* Baltimore: Johns Hopkins Press.

Fujita, Masahisa and Tomoya Mori (1996). "The Role of Ports in the Making of Major Cities: Self-Agglomeration and Hub-Effect." *Journal of Development Economics* 49, pp. 93–120.

Glaeser, Edward L. (1998). "Are Cities Dying?" *Journal of Economic Perspectives* 12, 2 (Spring), pp. 139–160.

Glaeser, Edward L., Heidi D. Kallal, José A. Scheinkman, and Andrei Shleifer (1992). "Growth in Cities." *Journal of Political Economy* 100, 6 (Dec.), pp. 1126-1152.

Glaeser, Edward L., José A. Scheinkman, and Andrei Shleifer (1995). "Economic Growth in a Cross-Section of Cities." *Journal of Monetary Economics* 36, pp. 117-143.

Gyourko, Joseph and Joseph Tracy (1989). "The Importance of Local Fiscal Conditions in Analyzing Local Labor Markets." *Journal of Political Economy* 97, 5 (Oct.), pp. 1208 – 1231.

Gyourko, Joseph and Joseph Tracy (1991). "The Structure of Local Public Finance and the Quality of Life." *Journal of Political Economy* 99, 4 (Aug.), pp. 774 – 806.

Harris, John R. and Michael P. Todaro (1970). "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review* 60, 1 (March), pp. 126–142.

Haurin, Donald R. (1980). "The Regional Distribution of Population, Migration, and Climate." *Quarterly Journal of Economics* 95, 2 (Sept.), pp. 293–308.

Henderson, Vernon (1974). "The Sizes and Types of Cities." *American Economic Review* 64, 4 (Sept.), pp. 640-656.

Henderson, Vernon (1988). *Urban Development: Theory, Fact, and Illusion.* New York: Oxford University Press.

Henderson, Vernon, Ari Kuncoro, Ari, and Matt Turner (1995). "Industrial Development in Cities." *Journal of Political Economy* 103, 5 (Oct.), pp. 1067-1085.

Islam, Nazrul (1995). "Growth Empirics: a Panel Data Approach." *Quarterly Journal of Economics* 110, 4 (November), pp. 1127–1170.

Jorgenson, Dale W. (1961). "The Development of a Dual Economy." *Economic Journal* 71, 282 (June), pp. 309–334.

Krugman, Paul (1991). *Geography and Trade.* Cambridge, MA: MIT Press.

Lewis, Arthur W. (1954). "Economic Development with Unlimited Supplies of Labour." *The Manchester School of Economics and Social Studies* 22, (May), pp. 139–191.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw (1994). "The Impact of Global Warning on Agriculture: a Ricardian Analysis." *American Economic Review* 84, 4 (Sept.), pp. 753–771.

Poor, Henry V. (1860). *History of the Railroads and Canals of the United States of America.* Reprint, New York: Augustus M. Kelley: 1970.

Quah, Danny (1993a). "Empirical Cross-section Dynamics in Economic Growth." *European Economic Review* 37, 2/3 (April), pp. 426–434.

Quah, Danny (1993b). "Galton's Fallacy and Tests of the Convergence Hypothesis." *Scandinavian Journal of Economics* 95, 4 (Dec.), pp. 427–443.

Quah, Danny (1996a). "Convergence Empirics Across Economies with (Some) Capital Mobility." *Journal of Economic Growth* 1, 1 (March), pp. 95–124.

Quah, Danny (1996b). "Empirics for Economic Growth and Convergence." *European Economic Review* 40, 6 (June), pp. 1353–1375.

Quigly, John M. (1998). "Urban Diversity and Economic Growth" *Journal of Economic Perspectives* 12, 2 (Spring), pp. 127–138.

Ranis, Gustav and John C. H. Fei (1961). "A Theory of Economic Development." *American Economic Review* 51, 4 (Sept.), pp. 533–565.

Rappaport, Jordan (2000). "Why Are Population Flows So Persistent?." Federal Reserve Bank of Kansas City Working Paper 99-13, (August).

Roback, Jennifer (1982). "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90, 6 (Dec.), pp. 1257 – 1278.

Rosen, Sherwin (1979). "Wage-Based Indexes of Urban Quality of Life." In Miezkowski and Straszheim, Eds., *Current Issues in Urban Economics.* Baltimore: Johns Hopkins University Press.

Sokoloff, Kenneth L. (1988). "Inventive Activity in Early Industrial America: Evidence from Patent Records, 1790-1846." *Journal of Economic History* 48, 4 (Dec.), pp. 813–850.

Southern Illinois University at Carbondale, Transportation Institute (1968). "A Study of River Ports and Terminals." Washington D.C.: U.S. Department of Commerce, National Bureau of Standards.

Stover, Mark Edward and Charles L. Leven (1992). "Methodological Issues in the Determination of the Quality of Life in Urban Areas." *Urban Studies* 29, 5, pp. 737 – 754.

Tanner, Henry S. (1840). *A Description of the Canals and Railroads of the United States.* Reprint, New York: Augustus M. Kelley: 1970.

Tiebout, Charles M. (1956). "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64, 5 (Oct.), pp. 416 – 424.

U.S. Department of Agriculture, Natural Resources Conservation Services. *1992 National Resources Inventory.* CD-Rom reissued May 1995. Fort Worth: Natural Resources Conservation Service.

U.S. Department of Commerce, Bureau of the Census. Decennial Census reports, 1890-1960 [investigator]. *Historical, Demographic, Economic, and Social Data: The United States, 1790-1970*, ICPSR Study No. 3 [computer file]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1992.

U.S. Deptartment of Commerce, Bureau of the Census. *Census of Population and Housing, 1970: Fourth Count Population Summary Tape, File C* [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1982. Cambridge, MA: National Bureau for Economic Research [distributor]. Also available as ICPSR Study No. 8107.

U.S. Deptartment of Commerce, Bureau of the Census. *Census of Population and Housing, 1970: Fourth Count Housing Summary Tape, File C* [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1982. Cambridge, MA: National Bureau for Economic Research [distributor]. Also available as ICPSR Study No. 8129.

U.S. Deptartment of Commerce, Bureau of the Census. *Census of Population and Housing, 1980: Summary Tape File 3C*, ICPSR Study No. 8038 [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1982. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1992.

U.S. Deptartment of Commerce, Bureau of the Census. *Census of Population and Housing, 1990: Summary Tape File 3C*, ICPSR Study No. 6054 [computer file]. Washington, DC: U.S. Department of Commerce, Bureau of the Census [producer], 1992. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1993.

U.S. Deptartment of Commerce, Bureau of the Census. "Census 2000 PHC-T-4. Ranking Tables for Counties: 1990 and 2000" [computer file]. Internet Release date: April 2, 2001. ¡http://blue.census.gov/population/cen2000/phc-t4/tab01.xls¿.

U.S. Deptartment of Commerce, Bureau of Economic Analysis. "CA05: Personal Income by Major Source and Earnings by Industry." [computer file]. *REIS Regional Economic Information System, 1969-98* CD-Rom (RCN-0250).

U.S. Deptartment of Commerce, Bureau of Economic Analysis. "CA25: Full-time and Part-time Employment by Industry." [computer file]. *REIS Regional Economic Information System, 1969-98* CD-Rom (RCN-0250).

U.S. Naval Oceanographic Office (1971). *World Port Index.*Publication 150, Fourth Edition. Washington: U.S. Government Printing Office.

Wright, Gavin (1990). "The Origins of American Industrial Success, 1879-1940." *American Economic Review* 80, 4 (Sept.), pp. 651-668.

Table 1: Distribution of Land, Population, Employment, and Income Across Continental U.S. Counties in 1998

	ALL (Continental)	"Coastal" Counties	"Inland" Counties	North Atlantic Coastal	South Atlantic Coastal	Gulf of Mexico Coastal	Pacific Ocean Coastal	Great Lakes Coastal	Navigable Rivers Coastal
Number of Counties	3,076	735	2,341	82	94	75	43	109	332
Percent of Continental U.S.:									
Land Area	100.0%	14.9%	85.1%	1.2%	1.5%	1.9%	2.0%	2.6%	5.6%
Population (2000)	100.0%	52.8%	47.2%	13.1%	4.8%	5.0%	10.8%	8.9%	10.2%
Civilian Employment	100.0%	54.3%	45.7%	13.3%	4.6%	4.9%	11.2%	9.1%	11.2%
Civilian Labor Income	100.0%	60.3%	39.7%	17.4%	4.4%	4.8%	12.9%	10.1%	10.8%
Capital Income	100.0%	59.6%	40.4%	16.0%	6.1%	5.1%	12.4%	9.8%	10.2%
Density Relative to Continental U.S.:									
Population (2000)	1.00	3.55	0.55	11.18	3.22	2.59	5.42	3.35	1.81
Civilian Employment	1.00	3.65	0.54	11.30	3.10	2.52	5.58	3.45	2.00
Civilian Labor Income	1.00	4.06	0.47	14.80	2.95	2.50	6.45	3.80	1.91
Capital Income	1.00	4.01	0.47	13.59	4.05	2.63	6.20	3.71	1.82
Income:									
Per Worker Labor Income	$33,071	$36,747	$28,702	$43,331	$31,436	$32,781	$38,276	$36,455	$31,613
Per Person Labor and Capital Income	$23,737	$27,052	$20,030	$30,728	$24,239	$23,281	$28,056	$26,641	$24,771
1990 Population by Age and Education:									
Age 0 to 17	25.6%	24.8%	26.6%	23.1%	23.9%	26.0%	25.0%	25.9%	25.5%
Age 18 to 64	61.8%	62.6%	60.9%	63.7%	62.5%	60.4%	64.5%	61.5%	61.4%
Age 65 and Older	12.6%	12.6%	12.6%	13.2%	13.6%	13.6%	10.5%	12.6%	13.1%
16 or more years of school (adults)	20.2%	22.2%	17.7%	25.0%	21.8%	18.8%	25.4%	19.6%	19.2%
Civilian Employed Persons by Industry:									
Natural Resources	3.5%	2.0%	5.2%	0.9%	1.9%	4.1%	2.3%	1.5%	2.7%
Manufacturing	12.5%	11.5%	13.6%	9.3%	7.2%	8.1%	12.2%	16.5%	12.8%
Finance, Insurance, Real Estate	7.7%	8.5%	6.7%	10.4%	8.1%	7.1%	8.5%	8.0%	7.4%
Services	31.3%	34.1%	28.0%	37.2%	35.5%	33.5%	35.3%	31.9%	30.9%

Ocean and Great Lake Coastal counties are counties with centroids within 50 kilometers of the respective coast; Navigable River Coastal are those counties not already classified as Ocean or Great Lake Coastal which have centroids within 25 kilometers of a river on which there was commercial navigation in 1968. For further information, see data appendix.

TABLE 2: Summary Statistics (All Continental Counties)

Variable	Obs	Mean	Std. Dev.	Min	Max
Land Area (km^2)	3,076	2,491	3,382	40	51,961
2000 Population	3,076	89,774	291,858	67	9,519,338
1998 Civilian Employment	3,076	50,462	177,409	124	5,291,228
1930 Population	3,064	39,745	135,681	195	3,982,123
1890 Population	2,702	22,886	55,627	3	1,515,301
2000 Population Density	3,076	86.5	585.1	0.0	20,915
1998 Civilian Employment Density	3,076	56.0	687.3	0.0	36,010
1930 Population Density	3,064	62.9	739.7	0.1	32,772
1890 Population Density	2,690	28.1	258.7	0.0	9,287
Log(1 + 2000 Population Density)	3,076	2.94	1.44	0.04	9.95
Log(1 + 1998 Civilian Employment Density)	3,076	2.32	1.38	0.05	10.49
Log(1 + 1930 Population Density)	3,064	2.54	1.16	0.08	10.40
Log(1 + 1890 Population Density)	2,690	2.28	1.08	0.00	9.14
Distance to Ocean Coast	3,076	638.0	483.3	0.4	1,875.0
Ocean Coast Dummy	294	1	0	1	1
Far Distance to Ocean Coast	2,782	653.3	462.1	0.1	1,825.0
Distance to Ocean Natural Harbor	3,076	660.5	468.0	2.6	1,887.8
Ocean Natural Harbor Dummy	108	1	0	1	1
Far Distance to Ocean Natural Harbor	2,968	633.5	460.4	0.2	1,837.8
Ocean Shoreline/km^2	3,076	0.01	0.03	0.00	0.66
Log(1 + Dist. to Ocean Coast)	3,076	5.93	1.34	0.34	7.54
Log(1 + Far Dist. to Ocean Coast)	2,782	2.33	0.86	0.00	3.61
Log(1 + Dist. to Ocean Natural Harbor)	3,076	6.13	1.01	1.27	7.54
Log(1 + Far Dist. to Ocean Natural Harbor)	2,968	2.30	0.85	0.00	3.61
Log(1 + Ocean Shoreline/km^2)	3,076	0.01	0.03	0.00	0.50
Distance to Great Lakes Coast	3,076	839.7	594.1	0.3	2,686.1
Great Lakes Dummy	109	1	0	1	1
Far Distance to Great Lakes Coast	2,967	819.8	583.4	0.6	2,636.1
Distance to Great Lakes Natural Harbor	3,076	928.4	566.6	5.0	2,703.2
Great Lakes Natural Harbor Dummy	15	1	0	1	1
Far Distance to Great Lakes Natural Harbor	3,061	882.8	564.5	2.1	2,653.2
Great Lakes Shoreline/km^2	3,076	0.00	0.01	0.00	0.18
Log(1 + Dist. to Great Lake Coast)	3,076	6.37	1.05	0.23	7.90
Log(1 + Far Dist. to Great Lake Coast)	2,967	2.57	0.81	0.01	3.96
Log(1 + Dist. to Great Lake Natural Harbor)	3,076	6.61	0.74	1.79	7.90
Log(1 + Far Dist. to Great Lake Natural Harbor)	3,061	2.70	0.70	0.04	3.97
Log(1 + Great Lake Shoreline/km^2)	3,076	0.00	0.01	0.00	0.17
Distance to Navigable River	3,076	248.7	244.7	1.2	1,246.0
Navigable River Dummy	371	1	0	1	1
Far Distance to Navigable River	2,705	256.2	243.7	0.2	1,221.0
Distance to Major River	3,076	74.1	67.1	0.5	436.5
Major River Dummy	828	1	0	1	1
Far Distance to Major River	2,248	71.8	65.0	0.1	411.5
Distance to 1890 Navigable River	3,076	101.0	131.6	0.1	770.8
1890 Navigable River Dummy	947	1	0	1	1
Far Distance to 1890 Navigable River	2,129	115.6	141.1	0.0	745.8
Log(1 + Dist. to Navigable River)	3,076	4.92	1.27	0.79	7.13
Log(1 + Far Dist. to Navigable River)	2,705	2.00	0.91	0.01	3.87
Log(1 + Dist. to Major River)	3,076	3.87	1.06	0.43	6.08
Log(1 + Far Dist. to Major River)	2,248	1.14	0.60	0.00	2.82
Log(1 + Dist. to 1890 Navigable River)	3,076	3.89	1.30	0.11	6.65
Log(1 + Far Dist. to 1890 Navigable River)	2,129	1.32	0.84	0.00	3.39

TABLE 3: Summary Statistics By Nearest Coast (1 of 3)

Variable	Obs	Mean	Std. Dev.	Min	Max
Counties for which North Atlantic is Closest Coast					
Land Area (km^2)	163	1,725	1,940	64	17,280
2000 Population	163	299,092	397,156	6,459	2,465,326
1998 Civilian Employment	163	169,404	272,929	2,629	2,646,630
1930 Population	161	176,508	343,100	3,678	2,560,401
1890 Population	159	87,832	169,402	3,268	1,515,301
2000 Population Density	163	692.4	2,326.6	1.7	20,915
1998 Civilian Employment Density	163	487.0	2,878.9	0.8	36,010
1930 Population Density	161	636.5	3,060.4	1.9	32,772
1890 Population Density	159	230.4	998.5	1.7	9,287
Log(1 + 2000 Population Density)	163	4.83	1.66	0.98	9.95
Log(1 + 1998 Civilian Employment Density)	163	4.19	1.68	0.61	10.49
Log(1 + 1930 Population Density)	161	4.09	1.67	1.05	10.40
Log(1 + 1890 Population Density)	159	3.67	1.33	1.00	9.14
Distance to Ocean Coast	163	70.5	62.5	0.4	214.2
Ocean Coast Dummy	82	1	0	1	1
Far Distance to Ocean Coast	81	73.8	44.9	0.3	164.2
Distance to Ocean Natural Harbor	163	98.0	65.0	2.6	351.1
Ocean Natural Harbor Dummy	49	1	0	1	1
Far Distance to Ocean Natural Harbor	114	78.2	53.9	0.2	301.1
Ocean Shoreline/km^2	163	0.04	0.08	0.00	0.66
Log(1 + Dist. to Ocean Coast)	163	3.69	1.24	0.34	5.37
Log(1 + Far Dist. to Ocean Coast)	81	0.82	0.40	0.01	1.44
Log(1 + Dist. to Ocean Natural Harbor)	163	4.30	0.90	1.27	5.86
Log(1 + Far Dist. to Ocean Natural Harbor)	114	0.85	0.41	0.00	1.93
Log(1 + Ocean Shoreline/km^2)	163	0.03	0.07	0.00	0.50
Counties for which South Atlantic is Closest Coast					
Land Area (km^2)	453	1,188	603	40	5,269
2000 Population	453	81,012	150,695	2,077	1,623,018
1998 Civilian Employment	453	44,867	97,545	404	847,559
1930 Population	449	25,807	28,882	2,466	318,587
1890 Population	402	16,599	11,985	2,863	103,394
2000 Population Density	453	94.4	262.3	2.4	3,243
1998 Civilian Employment Density	453	58.0	215.7	0.8	2,755
1930 Population Density	449	61.1	320.2	1.9	3,529
1890 Population Density	401	15.1	17.1	0.5	224
Log(1 + 2000 Population Density)	453	3.66	1.11	1.21	8.08
Log(1 + 1998 Civilian Employment Density)	453	2.88	1.23	0.59	7.92
Log(1 + 1930 Population Density)	449	2.97	0.89	1.08	8.17
Log(1 + 1890 Population Density)	401	2.58	0.57	0.44	5.42
Distance to Ocean Coast	453	207.2	156.3	1.1	598.3
Ocean Coast Dummy	94	1	0	1	1
Far Distance to Ocean Coast	359	206.4	138.3	0.4	548.3
Distance to Ocean Natural Harbor	453	249.8	146.3	5.7	617.1
Ocean Natural Harbor Dummy	24	1	0	1	1
Far Distance to Ocean Natural Harbor	429	212.2	140.2	0.9	567.1
Ocean Shoreline/km^2	453	0.01	0.05	0.00	0.37
Log(1 + Dist. to Ocean Coast)	453	4.83	1.27	0.73	6.40
Log(1 + Far Dist. to Ocean Coast)	359	1.45	0.63	0.01	2.46
Log(1 + Dist. to Ocean Natural Harbor)	453	5.29	0.78	1.91	6.43
Log(1 + Far Dist. to Ocean Natural Harbor)	429	1.48	0.59	0.02	2.49
Log(1 + Ocean Shoreline/km^2)	453	0.01	0.04	0.00	0.31

TABLE 3: Summary Statistics By Nearest Coast (2 of 3)

Variable	Obs	Mean	Std. Dev.	Min	Max
Counties for which Gulf of Mexico is Closest Coast					
Land Area (km^2)	815	2,232	1,692	334	17,164
2000 Population	815	62,138	185,685	67	3,400,578
1998 Civilian Employment	815	33,429	122,837	124	2,206,651
1930 Population	815	25,282	36,531	195	458,762
1890 Population	672	13,947	14,451	3	242,039
2000 Population Density	815	33.2	90.1	0.0	1,270
1998 Civilian Employment Density	815	17.9	58.4	0.1	793
1930 Population Density	815	14.4	34.5	0.1	904
1890 Population Density	672	8.5	19.1	0.0	474
Log(1 + 2000 Population Density)	815	2.65	1.19	0.04	7.15
Log(1 + 1998 Civilian Employment Density)	815	1.99	1.12	0.07	6.68
Log(1 + 1930 Population Density)	815	2.33	0.84	0.10	6.81
Log(1 + 1890 Population Density)	672	1.84	0.92	0.00	6.16
Distance to Ocean Coast	815	416.8	293.2	0.7	1,271.8
Ocean Coast Dummy	75	1	0	1	1
Far Distance to Ocean Coast	740	407.0	277.7	0.1	1,221.8
Distance to Ocean Natural Harbor	815	442.2	272.1	9.7	1,262.5
Ocean Natural Harbor Dummy	19	1	0	1	1
Far Distance to Ocean Natural Harbor	796	402.0	267.6	1.6	1,212.5
Ocean Shoreline/km^2	815	0.00	0.02	0.00	0.31
Log(1 + Dist. to Ocean Coast)	815	5.61	1.17	0.54	7.15
Log(1 + Far Dist. to Ocean Coast)	740	1.96	0.75	0.00	3.22
Log(1 + Dist. to Ocean Natural Harbor)	815	5.85	0.80	2.37	7.14
Log(1 + Far Dist. to Ocean Natural Harbor)	796	1.98	0.69	0.03	3.21
Log(1 + Ocean Shoreline/km^2)	815	0.00	0.02	0.00	0.27
Counties for which Pacific is Closest Coast					
Land Area (km^2)	344	7,386	7,414	121	51,961
2000 Population	344	168,775	625,399	493	9,519,338
1998 Civilian Employment	344	92,212	353,191	283	5,291,228
1930 Population	340	31,498	132,156	241	2,208,492
1890 Population	221	11,272	23,672	365	298,997
2000 Population Density	344	59.0	366.3	0.1	6,422
1998 Civilian Employment Density	344	40.6	336.2	0.0	6,093
1930 Population Density	340	24.6	316.9	0.1	5,832
1890 Population Density	221	14.2	165.1	0.0	2,456
Log(1 + 2000 Population Density)	344	2.15	1.60	0.10	8.77
Log(1 + 1998 Civilian Employment Density)	344	1.68	1.48	0.05	8.71
Log(1 + 1930 Population Density)	340	1.32	1.07	0.08	8.67
Log(1 + 1890 Population Density)	221	1.01	0.92	0.02	7.81
Distance to Ocean Coast	344	565.8	402.3	2.0	1,310.6
Ocean Coast Dummy	43	1	0	1	1
Far Distance to Ocean Coast	301	593.4	369.6	1.6	1,260.6
Distance to Ocean Natural Harbor	344	589.7	387.3	6.2	1,319.2
Ocean Natural Harbor Dummy	16	1	0	1	1
Far Distance to Ocean Natural Harbor	328	566.8	376.1	0.4	1,269.2
Ocean Shoreline/km^2	344	0.01	0.03	0.00	0.42
Log(1 + Dist. to Ocean Coast)	344	5.79	1.37	1.10	7.18
Log(1 + Far Dist. to Ocean Coast)	301	2.27	0.84	0.03	3.25
Log(1 + Dist. to Ocean Natural Harbor)	344	6.02	1.01	1.97	7.19
Log(1 + Far Dist. to Ocean Natural Harbor)	328	2.22	0.84	0.01	3.25
Log(1 + Ocean Shoreline/km^2)	344	0.01	0.03	0.00	0.35

TABLE 3: Summary Statistics By Nearest Coast (3 of 3)

Variable	Obs	Mean	Std. Dev.	Min	Max
Counties for which Great Lakes are Closest Coast					
Land Area (km^2)	1,301	1,907	1,669	160	20,452
2000 Population	1,301	63,024	201,946	444	5,376,741
1998 Civilian Employment	1,301	37,140	129,962	261	3,287,428
1930 Population	1,299	38,846	143,904	1,180	3,982,123
1890 Population	1,248	23,506	46,808	3	1,191,922
2000 Population Density	1,301	48.4	145.7	0.1	2,195
1998 Civilian Employment Density	1,301	29.3	106.8	0.1	1,864
1930 Population Density	1,299	32.8	175.0	0.3	5,203
1890 Population Density	1,237	19.5	85.8	0.0	2,860
Log(1 + 2000 Population Density)	1,301	2.84	1.32	0.10	7.69
Log(1 + 1998 Civilian Employment Density)	1,301	2.27	1.25	0.07	7.53
Log(1 + 1930 Population Density)	1,299	2.65	1.01	0.29	8.56
Log(1 + 1890 Population Density)	1,237	2.47	0.95	0.01	7.96
Distance to Great Lakes Coast	1,301	416.0	292.4	0.3	1,293.5
Great Lakes Dummy	109	1	0	1	1
Far Distance to Great Lakes Coast	1,192	402.0	278.9	0.6	1,243.5
Distance to Great Lakes Natural Harbor	1,301	492.5	267.2	5.0	1,310.5
Great Lakes Natural Harbor Dummy	15	1	0	1	1
Far Distance to Great Lakes Natural Harbor	1,286	447.8	264.1	2.1	1,260.5
Great Lakes Shoreline/km^2	1,301	0.00	0.01	0.00	0.18
Log(1 + Dist. to Great Lakes Coast)	1,301	5.63	1.11	0.23	7.17
Log(1 + Far Dist. to Great Lakes Coast)	1,192	1.95	0.75	0.01	3.23
Log(1 + Dist. to Great Lakes Natural Harbor)	1,301	6.02	0.66	1.79	7.18
Log(1 + Far Dist. to Great Lakes Natural Harbor)	1,286	2.12	0.61	0.04	3.25
Log(1 + Great Lakes Shoreline/km^2)	1,301	0.00	0.01	0.00	0.17

TABLE 4: Economic Density and Coastal Proximity

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Dependent Variable → **RHS Variables ↓**	**Log(1+2000 Population Density)**			**Log(1+1998 Employment Density)**		
Weather/Topography Controls	**Yes**	**No**	**Yes**	**Yes**	**No**	**Yes**
State Fixed Effects	**No**	**No**	**Yes**	**No**	**No**	**Yes**
Oceans:						
Ocean Coast Dummy	**0.82**	**0.69**	**0.74**	**0.87**	**0.85**	**0.79**
	(0.18)	**(0.20)**	**(0.17)**	**(0.19)**	**(0.20)**	**(0.17)**
Far Distance to Ocean Coast	**-0.25**	**-0.58**	**-0.25**	**-0.25**	**-0.46**	-0.19
	(0.09)	**(0.04)**	**(0.12)**	**(0.09)**	**(0.04)**	(0.13)
Great Lakes:						
Great Lakes Dummy	0.04	0.03	0.23	0.07	0.00	0.26
	(0.20)	(0.23)	(0.18)	(0.20)	(0.24)	(0.18)
Far Distance to Great Lakes	**-0.51**	**-0.56**	**-0.41**	**-0.42**	**-0.53**	**-0.36**
	(0.07)	**(0.05)**	**(0.16)**	**(0.07)**	**(0.05)**	**(0.16)**
Rivers:						
Navigable River Dummy	**0.32**	0.13	**0.23**	**0.36**	**0.21**	**0.28**
	(0.09)	(0.11)	**(0.09)**	**(0.10)**	**(0.11)**	**(0.09)**
Far Distance to Navigable River	-0.12	**-0.30**	**-0.21**	-0.09	**-0.23**	**-0.18**
	(0.06)	**(0.04)**	**(0.07)**	(0.06)	**(0.04)**	**(0.07)**
Observations	3,076	3,076	3,076	3,076	3,076	3,076
R^2	0.511	0.429	0.555	0.431	0.357	0.478
Adjusted R^2	0.507	0.428	0.545	0.427	0.356	0.466
Sum of Squared Residuals	3128.6	3655.5	2846.5	3371.0	3807.1	3090.8
Number of Indep. Variables	23	6	71	23	6	71

Standard errors in parenthesis are robust to spatial correlation using the Conley spatial estimator discussed in the text with a weighing that declines quadratically to zero for counties with centers 200 km apart. Bold type signifies coefficients significantly different from zero at the 0.05 level.

Coastal Proximity Measures: Ocean and Great Lakes dummies are one for counties with centroids within 50 km of respective coasts, zero otherwise. Ocean and Great Lakes "far distance" is measured by log(1+distance) - log(51) for counties with centroids more than 50 km from respective coast, zero otherwise. Distance measures to rivers on which there was navigation in 1968 are defined analogously with 25km as the boundary determining a positive dummy and log(26) as the adjustment factor on far distance.

TABLE 5: Coastal Proximity Controlling for History

Dependent Variable → **RHS Variables ↓**	**(1)** **ΔPop Density (1890-2000)**	**(2)** **2000 Pop Density**	**(3)** **2000 Pop Density**	**(4)** **ΔPop Density (1930-2000)**	**(5)** **2000 Pop Density**	**(6)** **2000 Pop Density**
Weather/Topography Controls	**Yes**	**No**	**Yes**	**Yes**	**No**	**Yes**
Historical Year	**-**	**1890**	**1890**	**-**	**1930**	**1930**
Historical Population Density:						
0 to 20 Percentile		0.99	0.90		1.01	1.11
		(0.11)	(0.10)		(0.08)	(0.06)
20 to 50 Percentile		**0.46**	**0.77**		0.89	**1.19**
		(0.13)	**(0.09)**		(0.10)	**(0.08)**
50 to 80 Percentile		**1.38**	**1.47**		**1.66**	**1.50**
		(0.18)	**(0.15)**		**(0.11)**	**(0.09)**
80 to 90 Percentile		**2.24**	**1.77**		0.90	**0.67**
		(0.35)	**(0.32)**		(0.17)	**(0.16)**
90 to 95 Percentile		1.24	1.37		**1.44**	**1.36**
		(0.38)	(0.36)		**(0.17)**	**(0.14)**
95 to 98 Percentile		**1.48**	**1.41**		**0.69**	**0.64**
		(0.15)	**(0.14)**		**(0.11)**	**(0.11)**
98 to 100 Percentile		**0.75**	**0.59**		**0.71**	**0.59**
		(0.07)	**(0.09)**		**(0.09)**	**(0.09)**
Oceans:						
Ocean Coast Dummy	**0.61**		**0.59**	**0.37**		**0.45**
	(0.11)		**(0.11)**	**(0.09)**		**(0.08)**
Far Distance to Ocean Coast	0.11		**0.16**	**0.15**		**0.17**
	(0.07)		**(0.07)**	**(0.06)**		**(0.07)**
Great Lakes:						
Great Lakes Dummy	0.16		0.24	-0.17		-0.09
	(0.14)		(0.14)	(0.09)		(0.10)
Far Distance to Great Lakes	**-0.24**		**-0.14**	**-0.23**		**-0.15**
	(0.05)		**(0.06)**	**(0.04)**		**(0.05)**
Rivers:						
Navigable River Dummy	0.10		0.03	0.06		0.07
	(0.06)		(0.06)	(0.05)		(0.05)
Far Distance to Navigable River	0.01		0.01	0.00		0.00
	(0.05)		(0.05)	(0.04)		(0.04)
Observations	2,690	2,690	2,690	3,064	3,064	3,064
R^2	0.331	0.591	0.719	0.330	0.746	0.834
Adjusted R^2	0.325	0.589	0.716	0.325	0.745	0.832
Sum of Squared Residuals	1518.7	2104.3	1444.6	1137.6	1610.0	1053.5
Number of Indep. Variables	23	7	30	23	7	30

Standard errors in parenthesis are robust to spatial correlation using the Conley spatial estimator discussed in the text with a weighing that declines quadratically to zero for counties with centers 200 km apart. Bold type signifies coefficients significantly different from zero at the 0.05 level for coastal proximity variables, significantly different from one at the 0.05 level for historical population density.

Coastal Proximity Measures: Ocean and Great Lakes dummies are one for counties with centroids within 50 km of respective coasts, zero otherwise. Ocean and Great Lakes "far distance" is measured by log(1+distance) - log(51) for counties with centroids more than 50 km from respective coast, zero otherwise. Distance measures to rivers on which there was navigation in 1968 are defined analogously with 25km as the boundary determining a positive dummy and log(26) as the adjustment factor on far distance.

Table 6: Coastal Proximity by Closest Coast

	(1)	(2)	(3)	(4)	(5)
Dependent Variable → **RHS Variables ↓**	**2000 Pop Density**	**ΔPop Density (1890-2000)**	**2000 Pop Density**	**ΔPop Density (1930-2000)**	**2000 Pop Density**
Weather/Topography Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Historical Population Controls	**No**	**No**	**Yes**	**No**	**Yes**
Historical Year	**-**	**-**	**1890**	**-**	**1930**
North Atlantic Counties:					
Ocean Coast Dummy	0.20	**-0.30**	-0.15	**-0.48**	**-0.31**
	(0.36)	**(0.14)**	(0.15)	**(0.14)**	**(0.13)**
Far Distance to Ocean Coast	**-1.10**	**-0.94**	**-0.94**	**-0.72**	**-0.86**
	(0.17)	**(0.10)**	**(0.11)**	**(0.12)**	**(0.12)**
South Atlantic Counties:					
Ocean Coast Dummy	**0.64**	**0.47**	**0.45**	**0.42**	**0.47**
	(0.29)	**(0.19)**	**(0.19)**	**(0.19)**	**(0.17)**
Far Distance to Ocean Coast	0.02	**0.33**	**0.27**	0.23	0.17
	(0.14)	**(0.10)**	**(0.10)**	(0.12)	(0.13)
Gulf of Mexico Counties:					
Ocean Coast Dummy	**0.87**	**1.14**	**0.95**	**0.76**	**0.69**
	(0.21)	**(0.18)**	**(0.18)**	**(0.14)**	**(0.14)**
Far Distance to Ocean Coast	-0.06	**0.34**	**0.31**	**0.22**	0.16
	(0.11)	**(0.10)**	**(0.10)**	**(0.09)**	(0.09)
Pacific Counties:					
Ocean Coast Dummy	**1.41**	**0.60**	0.45	0.20	-0.07
	(0.45)	**(0.30)**	(0.29)	(0.17)	(0.17)
Far Distance to Ocean Coast	0.35	-0.20	-0.14	-0.08	0.00
	(0.20)	(0.16)	(0.15)	(0.11)	(0.10)
Observations	3,076	2,690	2,690	3,064	3,064
R^2	0.540	0.361	0.740	0.358	0.846
Adjusted R^2	0.535	0.353	0.733	0.351	0.843
Sum of Squared Residuals	2945.7	1451.0	1337.6	1090.2	975.3
Number of Indep. Variables	33	33	68	33	68

Standard errors in parenthesis are robust to spatial correlation using the Conley spatial estimator discussed in the text with a weighing that declines quadratically to zero for counties with centers 200 km apart. Bold type signifies coefficients significantly different from zero at the 0.05 level.

Coastal Proximity Measures: Ocean and Great Lakes dummies are one for counties with centroids within 50 km of respective coasts or natural harbors, zero otherwise. Ocean and Great Lakes "far distance" is measured by log(1+distance) - log(51) for counties with centroids more than 50 km from respective coast or natural harbor, zero otherwise. Distance measures to rivers are defined analogously with 25km as the boundary determining a positive dummy and log(26) as the adjustment factor on far distance.

Table 7: Coastal Versus Harbor Proximity

Dependent Variable → RHS Variables ↓	**(1) 2000 Pop Density**	**(2) ΔPop Density (1890-2000)**	**(3) 2000 Pop Density**	**(4) ΔPop Density (1930-2000)**	**(5) 2000 Pop Density**
Weather/Topography Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Historical Population Controls	**No**	**No**	**Yes**	**No**	**Yes**
Historical Year	**-**	**-**	**1890**	**-**	**1930**
Oceans:					
Ocean Coast Dummy	0.35	**0.49**	**0.45**	**0.45**	**0.43**
	(0.20)	**(0.14)**	**(0.13)**	**(0.10)**	**(0.10)**
Far Distance to Ocean Coast	0.17	**0.31**	**0.28**	**0.28**	**0.26**
	(0.17)	**(0.11)**	**(0.12)**	**(0.10)**	**(0.11)**
Ocean Natural Harbor Dummy	**1.20**	**0.57**	**0.61**	0.00	0.21
	(0.23)	**(0.15)**	**(0.15)**	(0.12)	(0.11)
Far Distance to Natural Harbor	**-0.46**	-0.21	-0.14	-0.17	-0.13
	(0.17)	(0.12)	(0.12)	(0.09)	(0.09)
Shoreline/km	1.44	-0.87	-0.40	**-1.72**	-0.65
	(1.25)	(0.88)	(0.78)	**(0.85)**	(0.70)
Great Lakes:					
Great Lakes Dummy	0.07	0.26	**0.33**	0.01	0.07
	(0.20)	(0.16)	**(0.16)**	(0.10)	(0.11)
Far Distance to Great Lakes	-0.06	-0.04	0.02	0.05	0.08
	(0.11)	(0.09)	(0.09)	(0.07)	(0.08)
Great Lakes Natural Harbor Dummy	0.45	0.25	0.20	-0.11	-0.01
	(0.36)	(0.28)	(0.29)	(0.20)	(0.20)
Far Distance to Natural Harbor	**-0.71**	**-0.33**	**-0.29**	**-0.40**	**-0.35**
	(0.14)	**(0.11)**	**(0.12)**	**(0.09)**	**(0.09)**
Shoreline/km	-0.17	**-3.89**	-3.47	**-4.35**	**-4.06**
	(2.84)	**(1.91)**	(2.02)	**(1.96)**	**(1.82)**
Rivers:					
Navigable River Dummy	0.01	-0.06	-0.12	-0.07	-0.04
	(0.12)	(0.09)	(0.09)	(0.07)	(0.07)
Far Distance to Navigable River	-0.11	0.05	0.03	0.00	-0.01
	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
"Major" River Dummy	**0.24**	0.10	0.12	**0.11**	0.10
	(0.08)	(0.07)	(0.07)	**(0.05)**	(0.05)
Far Distance to "Major" River	-0.11	**-0.12**	**-0.10**	-0.05	-0.04
	(0.06)	**(0.05)**	**(0.05)**	(0.04)	(0.04)
Observations	3,076	2,690	2,690	3,064	3,064
R^2	0.558	0.363	0.730	0.351	0.838
Adjusted R^2	0.553	0.356	0.727	0.344	0.836
Sum of Squared Residuals	2829.4	1445.6	1385.1	1102.9	1025.1
Number of Indep. Variables	31	31	38	31	38

Standard errors in parenthesis are robust to spatial correlation using the Conley spatial estimator discussed in the text with a weighing that declines quadratically to zero for counties with centers 200 km apart. Bold type signifies coefficients significantly different from zero at the 0.05 level.

Coastal Proximity Measures: Ocean and Great Lakes dummies are one for counties with centroids within 50 km of respective coasts or natural harbors, zero otherwise. Ocean and Great Lakes "far distance" is measured by log(1+distance) - log(51) for counties with centroids more than 50 km from respective coast or natural harbor, zero otherwise. Distance measures to rivers are defined analogously with 25km as the boundary determining a positive dummy and log(26) as the adjustment factor on far distance.

Table 8: Coastal Versus Harbor Proximity By Closest Coast

	(1)	(2)	(3)	(4)	(5)
Dependent Variable → RHS Variables ↓	**2000 Pop Density**	**ΔPop Density (1890-2000)**	**2000 Pop Density**	**ΔPop Density (1930-2000)**	**2000 Pop Density**
Weather/Topography Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Historical Population Controls	**No**	**No**	**Yes**	**No**	**Yes**
Historical Year	**-**	**-**	**1890**	**-**	**1930**
North Atlantic Counties:					
Ocean Coast Dummy	-0.17	-0.31	-0.30	**-0.32**	**-0.33**
	(0.27)	(0.17)	(0.16)	**(0.15)**	**(0.13)**
Far Distance to Ocean Coast	0.03	**-0.58**	**-0.56**	**-0.62**	**-0.50**
	(0.36)	**(0.20)**	**(0.20)**	**(0.20)**	**(0.16)**
Ocean Natural Harbor Dummy	**0.95**	0.13	0.38	-0.18	0.25
	(0.30)	(0.19)	(0.20)	(0.22)	(0.17)
Far Distance to Natural Harbor	**-1.11**	-0.38	-0.41	-0.15	-0.41
	(0.35)	(0.25)	(0.27)	(0.19)	(0.21)
Shoreline/km^2	1.12	-0.25	0.88	-1.31	-0.34
	(2.05)	(0.80)	(0.81)	(1.00)	(0.58)
South Atlantic Counties:					
Ocean Coast Dummy	0.16	0.31	0.31	**0.54**	0.42
	(0.42)	(0.28)	(0.27)	**(0.20)**	(0.23)
Far Distance to Ocean Coast	0.03	0.15	0.03	0.06	0.00
	(0.31)	(0.19)	(0.19)	(0.19)	(0.23)
Ocean Natural Harbor Dummy	1.06	0.62	**0.68**	-0.01	0.35
	(0.60)	(0.36)	**(0.32)**	(0.27)	(0.29)
Far Distance to Natural Harbor	0.07	0.27	0.33	0.24	0.25
	(0.41)	(0.26)	(0.25)	(0.20)	(0.22)
Shoreline/km^2	2.59	0.46	-0.49	-2.44	-0.91
	(2.90)	(1.68)	(1.67)	(1.40)	(1.89)
Gulf of Mexico Counties:					
Ocean Coast Dummy	**0.58**	**1.10**	**0.85**	**0.71**	**0.70**
	(0.19)	**(0.23)**	**(0.23)**	**(0.15)**	**(0.15)**
Far Distance to Ocean Coast	0.26	**0.56**	**0.55**	**0.42**	**0.43**
	(0.23)	**(0.19)**	**(0.18)**	**(0.13)**	**(0.14)**
Ocean Natural Harbor Dummy	**0.93**	**0.80**	**0.91**	0.16	-0.04
	(0.30)	**(0.27)**	**(0.26)**	(0.16)	(0.19)
Far Distance to Natural Harbor	-0.34	-0.27	-0.27	-0.27	**-0.33**
	(0.24)	(0.20)	(0.19)	(0.14)	**(0.14)**
Shoreline/km^2	0.88	-3.29	-2.43	0.09	0.08
	(2.08)	(2.20)	(2.32)	(1.30)	(1.45)
Pacific Counties:					
Ocean Coast Dummy	**1.05**	**0.53**	0.32	0.17	-0.20
	(0.39)	**(0.25)**	(0.24)	(0.18)	(0.16)
Far Distance to Ocean Coast	0.49	-0.09	-0.03	-0.07	-0.09
	(0.34)	(0.28)	(0.26)	(0.17)	(0.15)
Ocean Natural Harbor Dummy	**1.60**	**0.81**	**0.96**	0.24	0.31
	(0.39)	**(0.26)**	**(0.25)**	(0.23)	(0.17)
Far Distance to Natural Harbor	-0.11	-0.11	-0.05	0.01	0.16
	(0.32)	(0.25)	(0.24)	(0.16)	(0.14)
Shoreline/km^2	**5.38**	0.62	**2.93**	1.32	**3.56**
	(2.11)	(3.81)	**(1.42)**	(2.87)	**(0.96)**
Observations	3,076	2,690	2,690	3,064	3,064
R^2	0.578	0.387	0.751	0.374	0.850
Adjusted R^2	0.571	0.375	0.743	0.364	0.845
Sum of Squared Residuals	2700.0	1392.7	1281.1	1062.4	951.8
Number of Indep. Variables	50	50	85	50	85

Regressions are analogous to those reported in corresponding column of Table 7. All control for Great Lakes, navigable river, and major river proximity as there. Standard errors are robust to spatial correlation. Bold type signifies coefficients significantly different from zero at the 0.05 level

Table 9: Robustness to Dummy/Far Distance Boundary

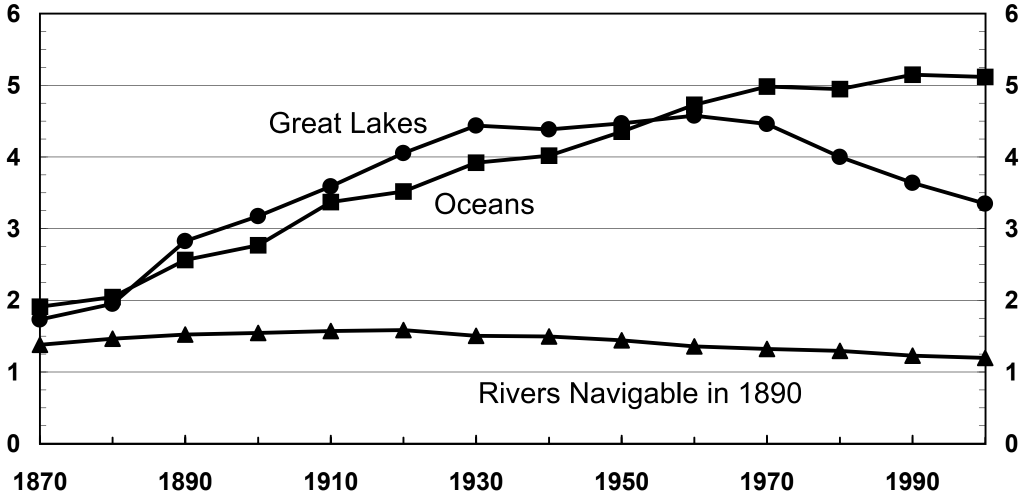
Dependent Variable → RHS Variables ↓	(1) 2000 Pop Density	(2) ΔPop Density (1890-2000)	(3a) 2000 Pop Density	(3b) 2000 Pop Density	(4) ΔPop Density (1930-2000)	(5a) 2000 Pop Density	(5b) 2000 Pop Density
Weather/Topography Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Historical Population Controls	**No**	**No**	**Yes**	**Yes**	**No**	**Yes**	**Yes**
Historical Adjacency Controls	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
Historical Year	**-**	**-**	**1890**	**1890**	**-**	**1930**	**1930**
Observations	3,076	2,690	2,690	2,677	3,064	3,064	3,060
# of Indep. Variables	23	23	30	36	23	30	36
A. 25 km/25 km/15 km Boundary (Ocean/Great Lakes/Navigable Rivers)							
Ocean Coast Dummy	0.31	**0.31**	**0.34**	**0.40**	0.14	**0.24**	**0.26**
	(0.20)	**(0.12)**	**(0.12)**	**(0.12)**	(0.10)	**(0.09)**	**(0.09)**
Far Distance to Ocean Coast	**-0.37**	0.00	0.06	0.12	0.07	0.08	0.12
	(0.10)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)
Great Lakes Dummy	-0.13	-0.09	0.01	0.10	**-0.35**	**-0.25**	-0.13
	(0.20)	(0.14)	(0.14)	(0.12)	**(0.10)**	**(0.10)**	(0.09)
Far Distance to Great Lakes	**-0.46**	**-0.24**	**-0.16**	**-0.13**	**-0.20**	**-0.13**	-0.08
	(0.07)	**(0.05)**	**(0.05)**	**(0.04)**	**(0.04)**	**(0.04)**	(0.04)
Navigable River Dummy	**0.35**	0.09	0.01	0.04	-0.01	-0.01	0.04
	(0.11)	(0.08)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)
Far Distance to Navigable River	**-0.12**	0.01	0.01	0.07	-0.01	-0.01	0.04
	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
R^2	0.507	0.319	0.714	0.744	0.321	0.831	0.848
B. 50 km/50 km/25 km Boundary (Ocean/Great Lakes/Navigable Rivers)							
Ocean Coast Dummy	**0.82**	**0.61**	**0.59**	**0.56**	**0.37**	**0.45**	**0.40**
	(0.18)	**(0.11)**	**(0.11)**	**(0.11)**	**(0.09)**	**(0.08)**	**(0.08)**
Far Distance to Ocean Coast	**-0.25**	0.11	**0.16**	**0.22**	**0.15**	**0.17**	**0.20**
	(0.09)	(0.07)	**(0.07)**	**(0.07)**	**(0.06)**	**(0.07)**	**(0.07)**
Great Lakes Dummy	0.04	0.16	0.24	0.17	-0.17	-0.09	-0.10
	(0.20)	(0.14)	(0.14)	(0.12)	(0.09)	(0.10)	(0.08)
Far Distance to Great Lakes	**-0.51**	**-0.24**	**-0.14**	**-0.13**	**-0.23**	**-0.15**	**-0.11**
	(0.07)	**(0.05)**	**(0.06)**	**(0.05)**	**(0.04)**	**(0.05)**	**(0.05)**
Navigable River Dummy	**0.32**	0.10	0.03	0.00	0.06	0.07	0.06
	(0.09)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
Far Distance to Navigable River	-0.12	0.01	0.01	0.05	0.00	0.00	0.05
	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
R^2	0.511	0.331	0.719	0.748	0.330	0.834	0.850
C. 75 km/75 km/35 km Boundary (Ocean/Great Lakes/Navigable Rivers)							
Ocean Coast Dummy	**0.86**	**0.75**	**0.69**	**0.59**	**0.41**	**0.48**	**0.37**
	(0.18)	**(0.12)**	**(0.12)**	**(0.11)**	**(0.09)**	**(0.09)**	**(0.09)**
Far Distance to Ocean Coast	-0.16	**0.23**	**0.27**	**0.32**	**0.21**	**0.24**	**0.25**
	(0.10)	**(0.07)**	**(0.07)**	**(0.07)**	**(0.06)**	**(0.06)**	**(0.06)**
Great Lakes Dummy	-0.05	0.16	0.24	0.16	-0.08	-0.04	-0.09
	(0.19)	(0.14)	(0.14)	(0.11)	(0.10)	(0.11)	(0.09)
Far Distance to Great Lakes	**-0.58**	**-0.27**	**-0.15**	**-0.15**	**-0.26**	**-0.17**	**-0.14**
	(0.08)	**(0.06)**	**(0.06)**	**(0.06)**	**(0.05)**	**(0.05)**	**(0.05)**
Navigable River Dummy	**0.29**	0.08	0.01	-0.05	0.03	0.05	0.01
	(0.10)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
Far Distance to Navigable River	-0.11	0.00	-0.01	0.03	0.00	-0.01	0.04
	(0.07)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
R^2	0.507	0.340	0.722	0.750	0.334	0.835	0.851
D. 100 km/100 km/50 km Boundary (Ocean/Great Lakes/Navigable Rivers)							
Ocean Coast Dummy	**0.74**	**0.65**	**0.59**	**0.49**	**0.31**	**0.37**	**0.28**
	(0.17)	**(0.11)**	**(0.11)**	**(0.11)**	**(0.09)**	**(0.09)**	**(0.09)**
Far Distance to Ocean Coast	-0.10	**0.31**	**0.35**	**0.40**	**0.25**	**0.27**	**0.29**
	(0.10)	**(0.07)**	**(0.07)**	**(0.07)**	**(0.06)**	**(0.06)**	**(0.06)**
Great Lakes Dummy	-0.21	0.05	0.13	0.09	-0.10	-0.07	-0.09
	(0.19)	(0.13)	(0.14)	(0.11)	(0.10)	(0.10)	(0.09)
Far Distance to Great Lakes	**-0.68**	**-0.33**	**-0.18**	**-0.18**	**-0.29**	**-0.20**	**-0.16**
	(0.09)	**(0.07)**	**(0.08)**	**(0.07)**	**(0.06)**	**(0.06)**	**(0.06)**
Navigable River Dummy	**0.26**	0.06	0.00	-0.08	0.03	0.03	-0.03
	(0.10)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
Far Distance to Navigable River	-0.09	0.00	0.00	0.02	0.00	-0.01	0.03
	(0.07)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
R^2	0.500	0.335	0.721	0.749	0.332	0.834	0.851

Table 10: Robustness to Alternate Harbor Measures

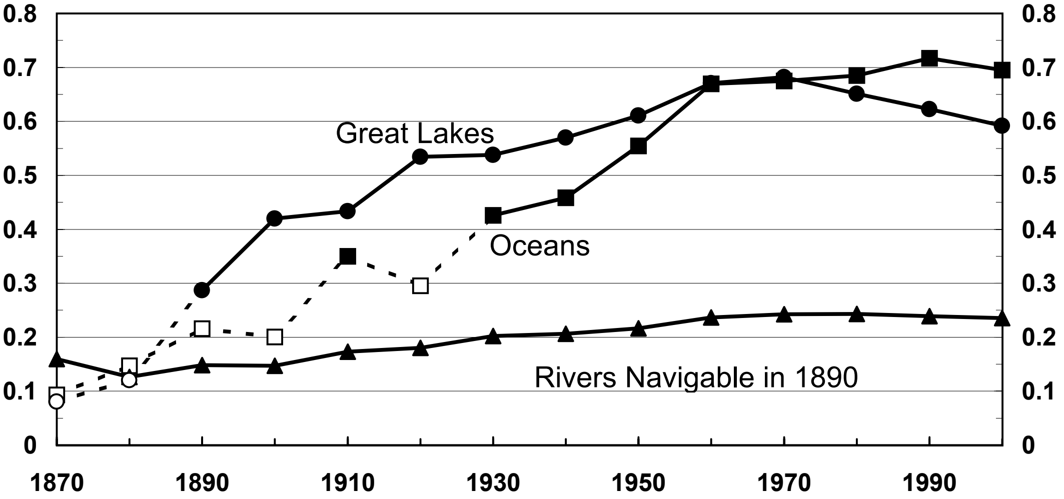
Dependent Variable → **RHS Variables ↓**	**(1)** **2000 Pop Density**	**(2)** **ΔPop Density (1890-2000)**	**(3a)** **2000 Pop Density**	**(3b)** **2000 Pop Density**	**(4)** **ΔPop Density (1930-2000)**	**(5a)** **2000 Pop Density**	**(5b)** **2000 Pop Density**
Weather/Topography Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Navigable River Proximity Controls	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Historical Population Controls	**No**	**No**	**Yes**	**Yes**	**No**	**Yes**	**Yes**
Historical Adjacency Controls	**No**	**No**	**No**	**Yes**	**No**	**No**	**Yes**
Historical Year	**-**	**-**	**1890**	**1890**	**-**	**1930**	**1930**
Observations	3,076	2,690	2,690	2,677	3,064	3,064	3,060
# of Indep. Variables	27	27	34	40	27	34	40
A. Medium/Large Natural Seaport:							
Ocean Coast Dummy	**0.39**	**0.41**	**0.40**	**0.44**	**0.34**	**0.38**	**0.39**
	(0.18)	**(0.13)**	**(0.12)**	**(0.12)**	**(0.10)**	**(0.10)**	**(0.10)**
Far Distance to Ocean Coast	0.18	**0.31**	**0.28**	**0.30**	**0.27**	**0.25**	0.21
	(0.17)	**(0.12)**	**(0.12)**	**(0.12)**	**(0.11)**	**(0.11)**	(0.12)
Ocean Harbor Dummy	**1.20**	**0.57**	**0.61**	**0.37**	-0.01	0.21	-0.03
	(0.24)	**(0.15)**	**(0.15)**	**(0.16)**	(0.13)	(0.12)	(0.11)
Far Distance to Harbor	**-0.49**	-0.23	-0.16	-0.12	-0.17	-0.13	-0.06
	(0.17)	(0.12)	(0.12)	(0.12)	(0.09)	(0.09)	(0.10)
Great Lakes Dummy	0.02	0.13	0.22	0.17	-0.12	-0.06	-0.04
	(0.19)	(0.15)	(0.15)	(0.13)	(0.09)	(0.10)	(0.09)
Far Distance to Great Lakes	-0.05	-0.05	0.02	0.03	0.03	0.07	0.11
	(0.11)	(0.08)	(0.09)	(0.08)	(0.07)	(0.08)	(0.08)
Great Lakes Harbor Dummy	0.48	0.28	0.22	0.06	-0.08	0.01	-0.19
	(0.36)	(0.29)	(0.30)	(0.27)	(0.21)	(0.21)	(0.16)
Far Distance to Great Lakes Harbor	**-0.68**	**-0.30**	**-0.26**	**-0.27**	**-0.37**	**-0.33**	**-0.32**
	(0.14)	**(0.11)**	**(0.12)**	**(0.11)**	**(0.09)**	**(0.09)**	**(0.09)**
R^2	0.550	0.353	0.727	0.751	0.340	0.837	0.852
B. Small/Medium/Large Natural Seaport:							
Ocean Coast Dummy	0.26	**0.32**	**0.34**	**0.39**	**0.27**	**0.33**	**0.35**
	(0.16)	**(0.12)**	**(0.12)**	**(0.11)**	**(0.09)**	**(0.09)**	**(0.09)**
Far Distance to Ocean Coast	**-0.40**	0.07	0.14	**0.23**	**0.24**	**0.28**	**0.30**
	(0.12)	(0.09)	(0.09)	**(0.09)**	**(0.07)**	**(0.08)**	**(0.08)**
Ocean Harbor Dummy	**1.28**	**0.65**	**0.59**	**0.42**	**0.21**	**0.30**	0.13
	(0.20)	**(0.13)**	**(0.13)**	**(0.12)**	**(0.10)**	**(0.11)**	(0.10)
Far Distance to Harbor	**0.34**	0.11	0.06	0.01	-0.10	-0.13	-0.13
	(0.12)	(0.09)	(0.09)	(0.09)	(0.07)	(0.07)	(0.07)
Great Lakes Dummy	0.07	0.20	0.27	0.19	**-0.25**	-0.17	-0.16
	(0.26)	(0.20)	(0.19)	(0.17)	**(0.12)**	(0.12)	(0.11)
Far Distance to Great Lakes	0.50	-0.14	-0.21	-0.21	-0.21	-0.23	-0.15
	(0.31)	(0.21)	(0.20)	(0.17)	(0.16)	(0.16)	(0.14)
Great Lakes Harbor Dummy	**-1.05**	**-0.60**	**-0.54**	**-0.40**	-0.12	-0.20	-0.08
	(0.36)	**(0.22)**	**(0.21)**	**(0.18)**	(0.15)	(0.16)	(0.14)
Far Distance to Great Lakes Harbor	**-1.34**	-0.20	0.01	0.06	0.05	0.17	0.14
	(0.33)	(0.23)	(0.23)	(0.19)	(0.17)	(0.18)	(0.15)
R^2	0.537	0.345	0.724	0.750	0.335	0.836	0.851
C. Very Small/Small/Medium/Large Natural Seaport:							
Ocean Coast Dummy	0.34	**0.27**	**0.28**	**0.38**	**0.20**	**0.28**	**0.28**
	(0.18)	**(0.12)**	**(0.12)**	**(0.10)**	**(0.10)**	**(0.09)**	**(0.08)**
Far Distance to Ocean Coast	**-0.51**	0.04	0.15	**0.24**	**0.24**	**0.28**	**0.31**
	(0.13)	(0.10)	(0.10)	**(0.10)**	**(0.08)**	**(0.09)**	**(0.09)**
Ocean Harbor Dummy	**0.73**	**0.49**	**0.44**	**0.26**	**0.24**	**0.24**	**0.15**
	(0.17)	**(0.12)**	**(0.11)**	**(0.09)**	**(0.08)**	**(0.09)**	**(0.08)**
Far Distance to Harbor	**0.41**	0.14	0.06	0.01	-0.09	-0.12	-0.13
	(0.13)	(0.10)	(0.10)	(0.09)	(0.07)	(0.07)	(0.07)
Great Lakes Dummy	0.18	0.34	0.40	0.25	-0.21	-0.12	-0.17
	(0.32)	(0.24)	(0.23)	(0.19)	(0.14)	(0.14)	(0.12)
Far Distance to Great Lakes	0.26	-0.31	-0.32	-0.28	-0.33	-0.34	-0.19
	(0.35)	(0.24)	(0.23)	(0.19)	(0.18)	(0.18)	(0.15)
Great Lakes Harbor Dummy	**-0.75**	**-0.64**	**-0.59**	-0.34	-0.21	-0.24	-0.11
	(0.36)	**(0.24)**	**(0.22)**	(0.19)	(0.15)	(0.15)	(0.13)
Far Distance to Great Lakes Harbor	**-1.11**	-0.05	0.13	0.13	0.16	0.27	0.17
	(0.38)	(0.26)	(0.26)	(0.21)	(0.19)	(0.20)	(0.17)
R^2	0.520	0.338	0.721	0.749	0.334	0.835	0.851

Figure 2: Coastal Concentration of U.S. Population

A. Population Density Relative to Continental U.S.

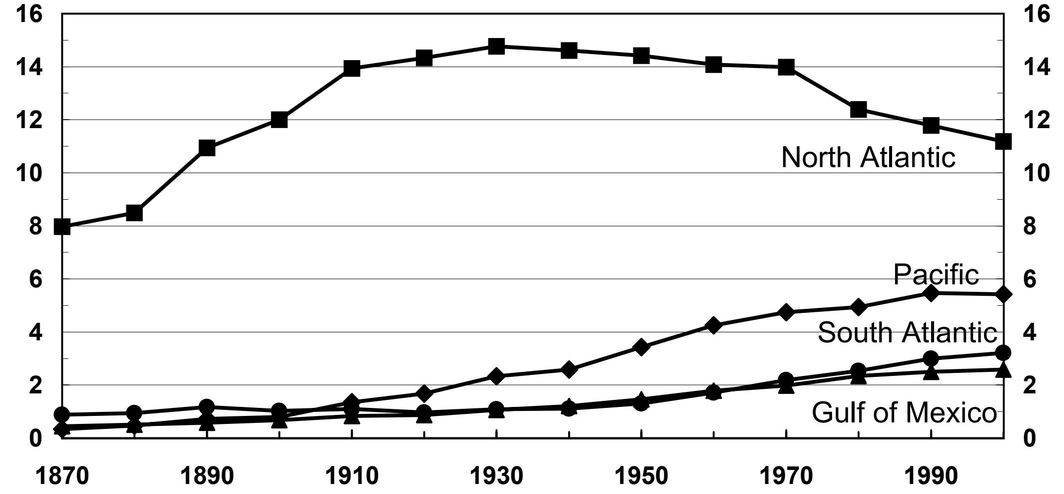


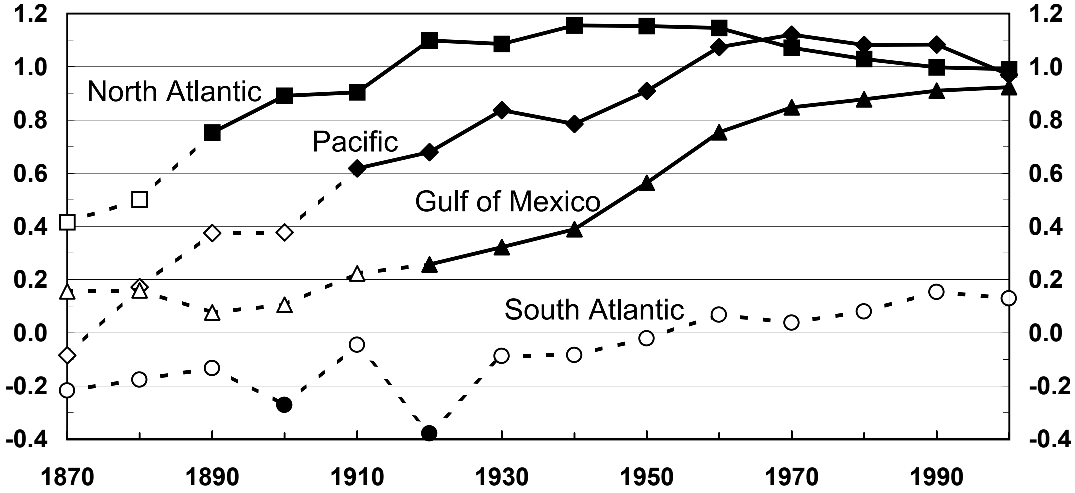
B. Partial Correlation Coefficients of Pop. Density



The ocean and Great Lake coastal categories are made up of those counties with centroids within 50 km of the respective coast; the navigable river coastal category is made up of those counties not already in the ocean or Great Lake coastal categories with centroids within 25 km of a river navigable in 1890 according to Fogel (1964). Panel A shows the aggregate population density of each of the categories relative to that of the continental United States in the same year. Panel B reports coefficients on category dummy variables from regressing log(1+Population Density) on these along with weather and topography variables as enumerated in the text along with state fixed effects. Open points (connected by dashed lines) represent coefficients not significant at the 0.05 level (using standard errors robust to spatial correlation as described in the text).

Figure 3: U.S. Population Density by Ocean Coast


A. Population Density Relative to Continental U.S.


B. Partial Correlation Coefficients of Pop. Density

Categories are made up of those counties with centroids within 50 km of the respective coast; "South Atlantic" coastal is composed of Atlantic coastal counties in Virginia, North Carolina, South Carolina, Georgia, and Florida. "North Atlantic" coastal is composed of the remaining Atlantic coastal counties. Panel A shows the aggregate population density of each of the categories relative to that of the continental United States in the same year. Panel B reports coefficients from regressing log(1+Population Density) on coastal category dummy variables while controlling for Great Lakes and 1890 navigable river proximity, weather and topography as enumerated in the text, and state fixed effects. Open points (connected by dashed lines) represent coefficients not significant at the 0.05 level (using standard errors robust to spatial correlation as described in the text).

Figure 4: Population Density Versus Coastal Distance

Graphical Representation of Partial Correlations Reported in Table 4 Column 1

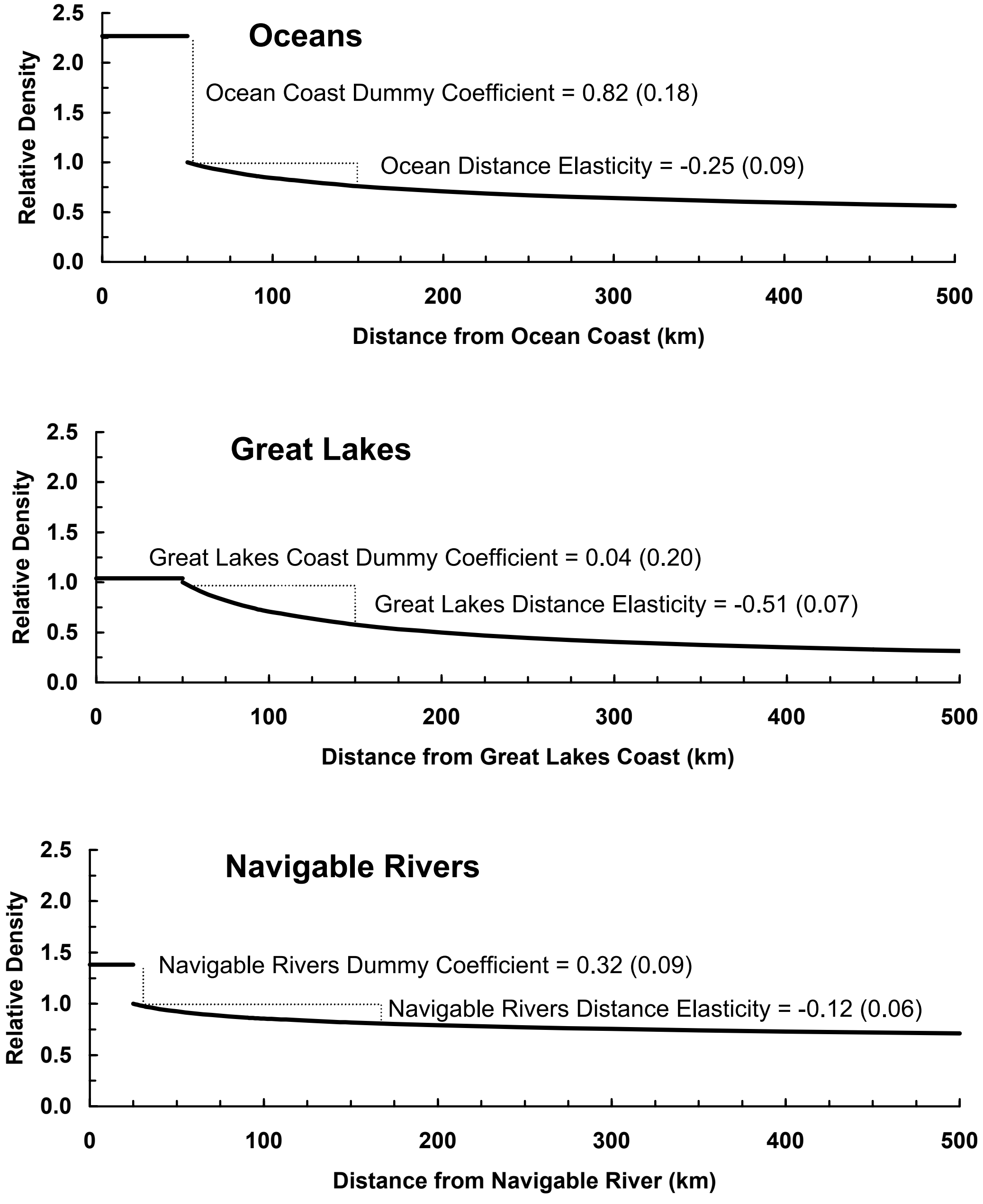
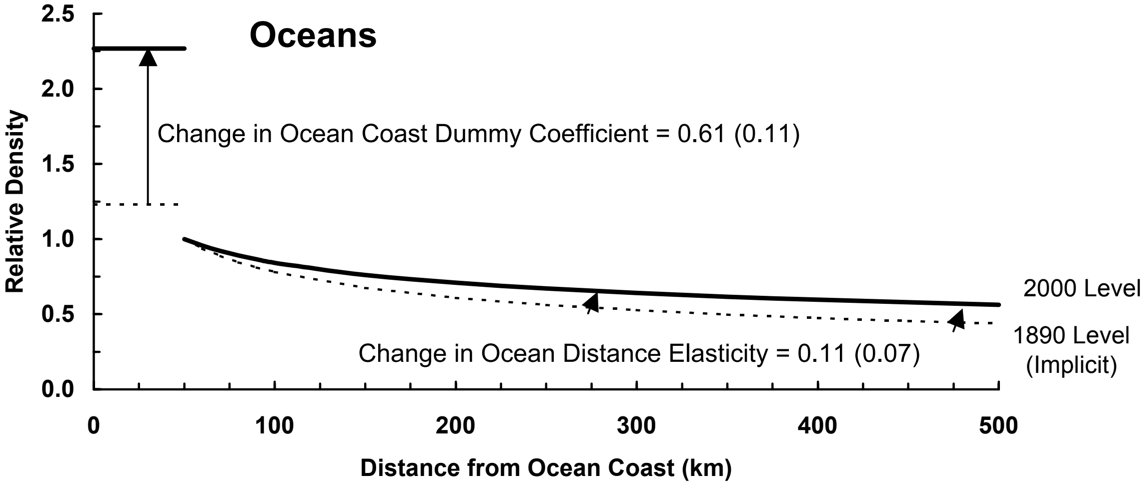


Figure 5: Change in Population Density 1890 to 2000 Versus Coastal Distance

Graphical Representation of Partial Correlations Reported in Table 5 Column 1
Combined with Partial Correlations Reported in Table 4 Column 1



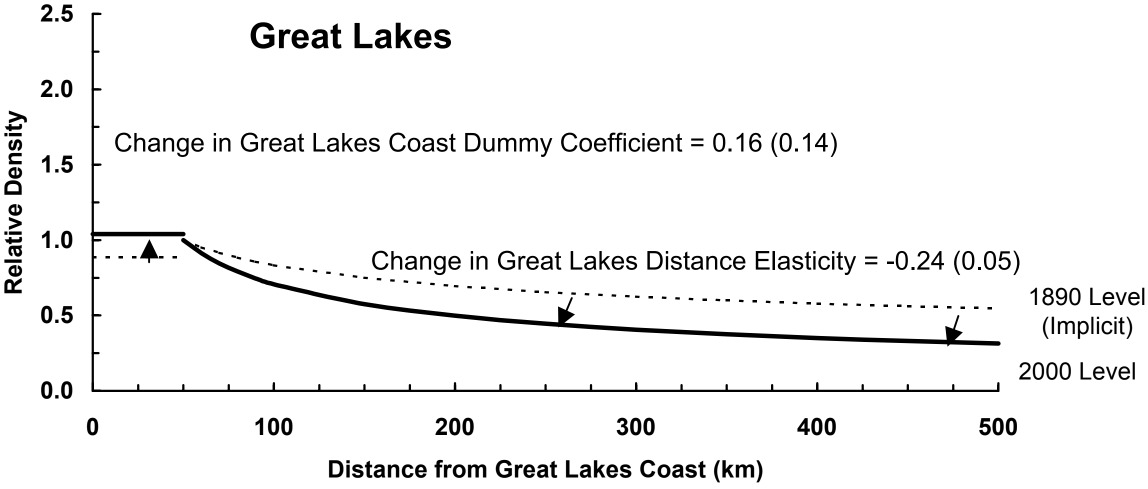
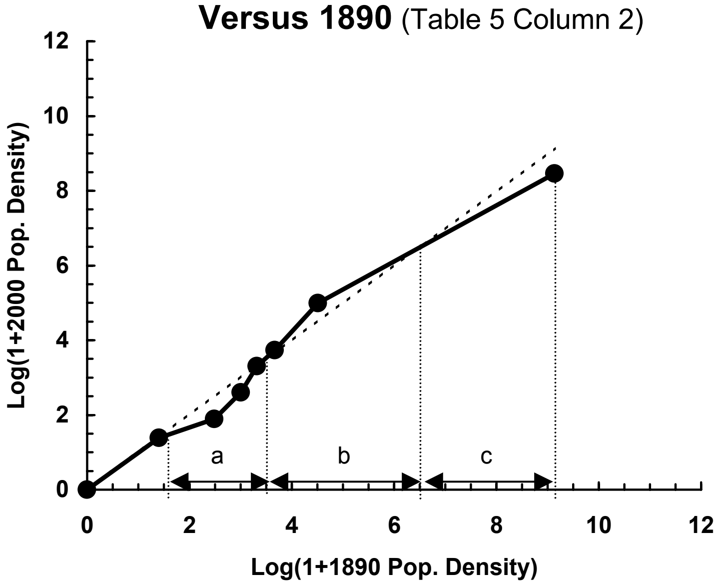


Figure 6: Current Versus Historical Population Density



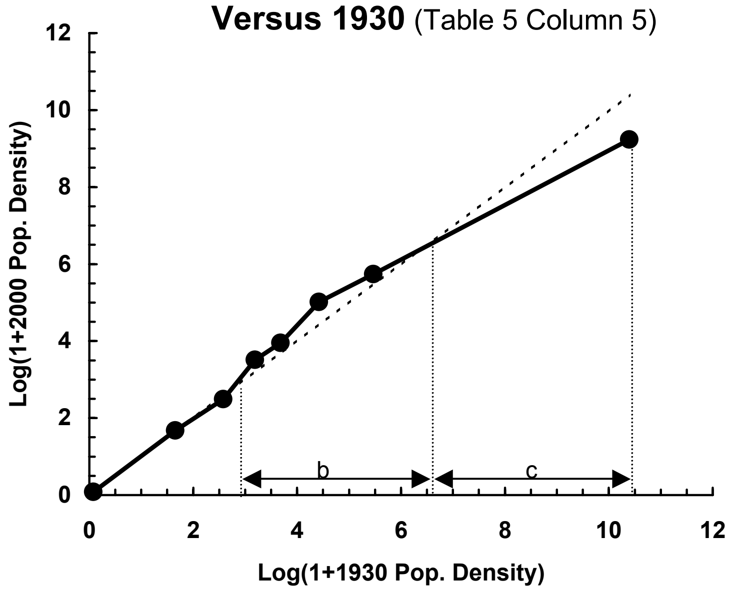
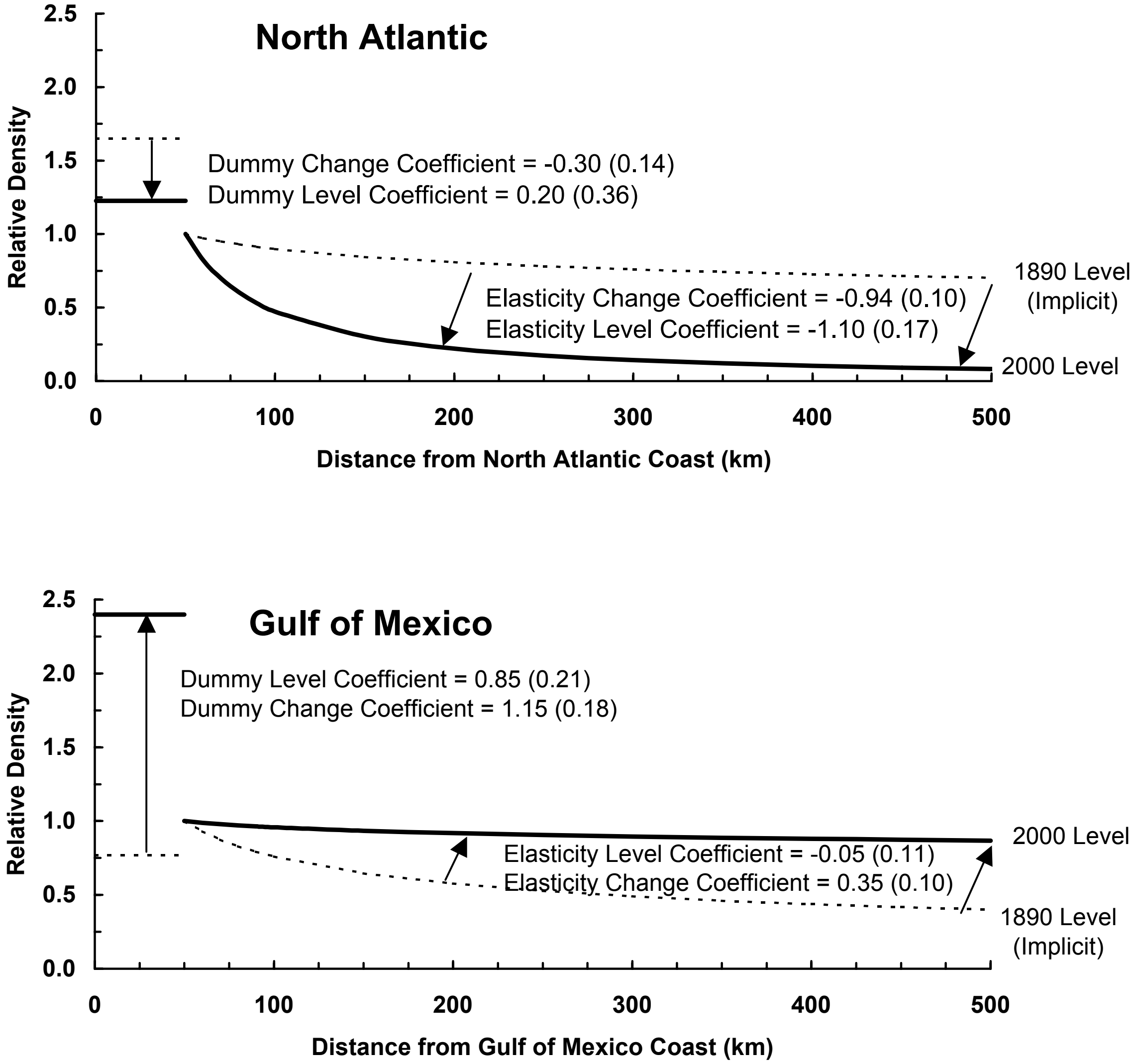
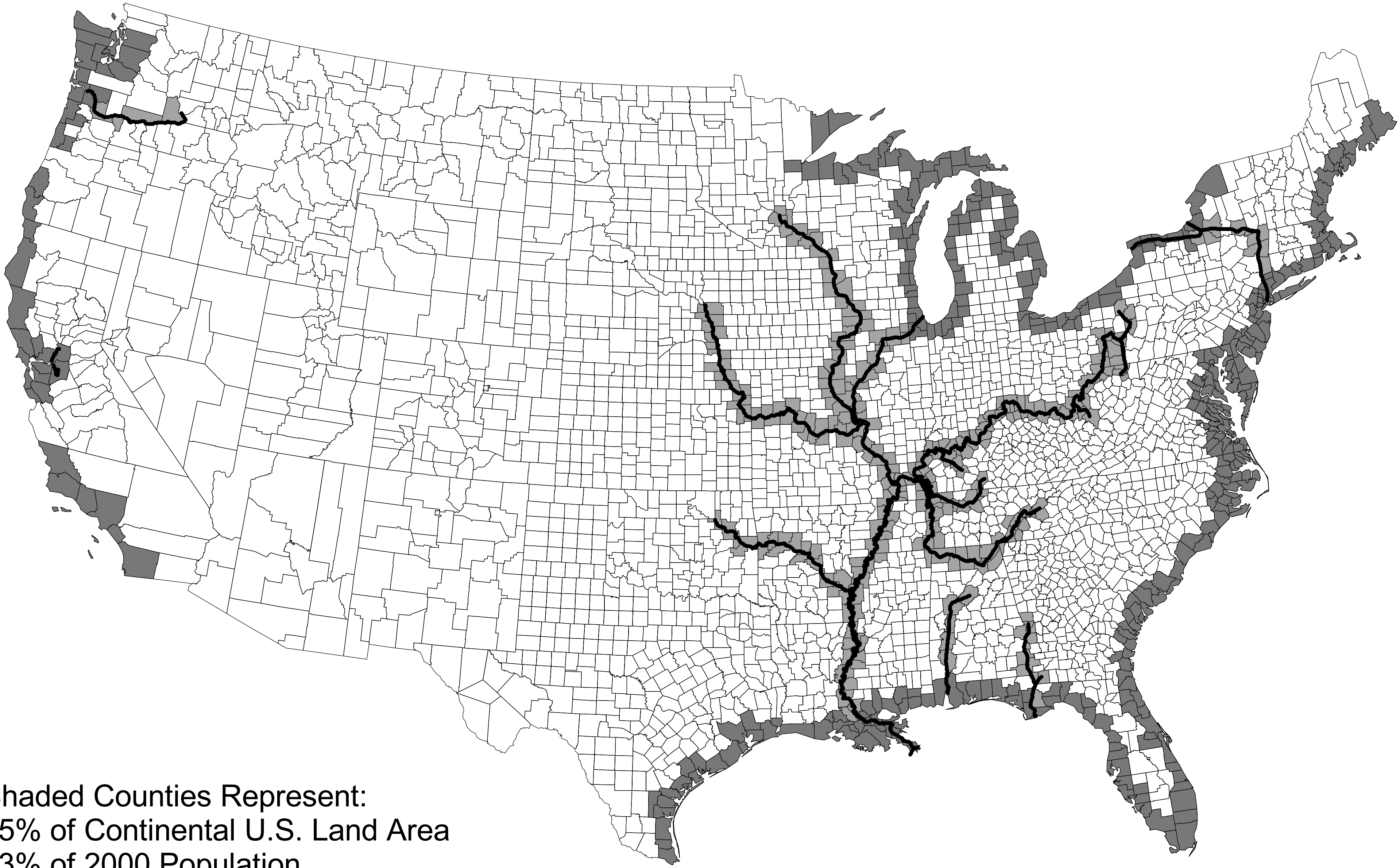


Figure 7: Change in Population Density 1890 to 2000 By Closest Coast

Graphical Representation of Partial Correlations Reported in Table 6 Column 2
Combined with Partial Correlations Reported in Table 6 Column 1



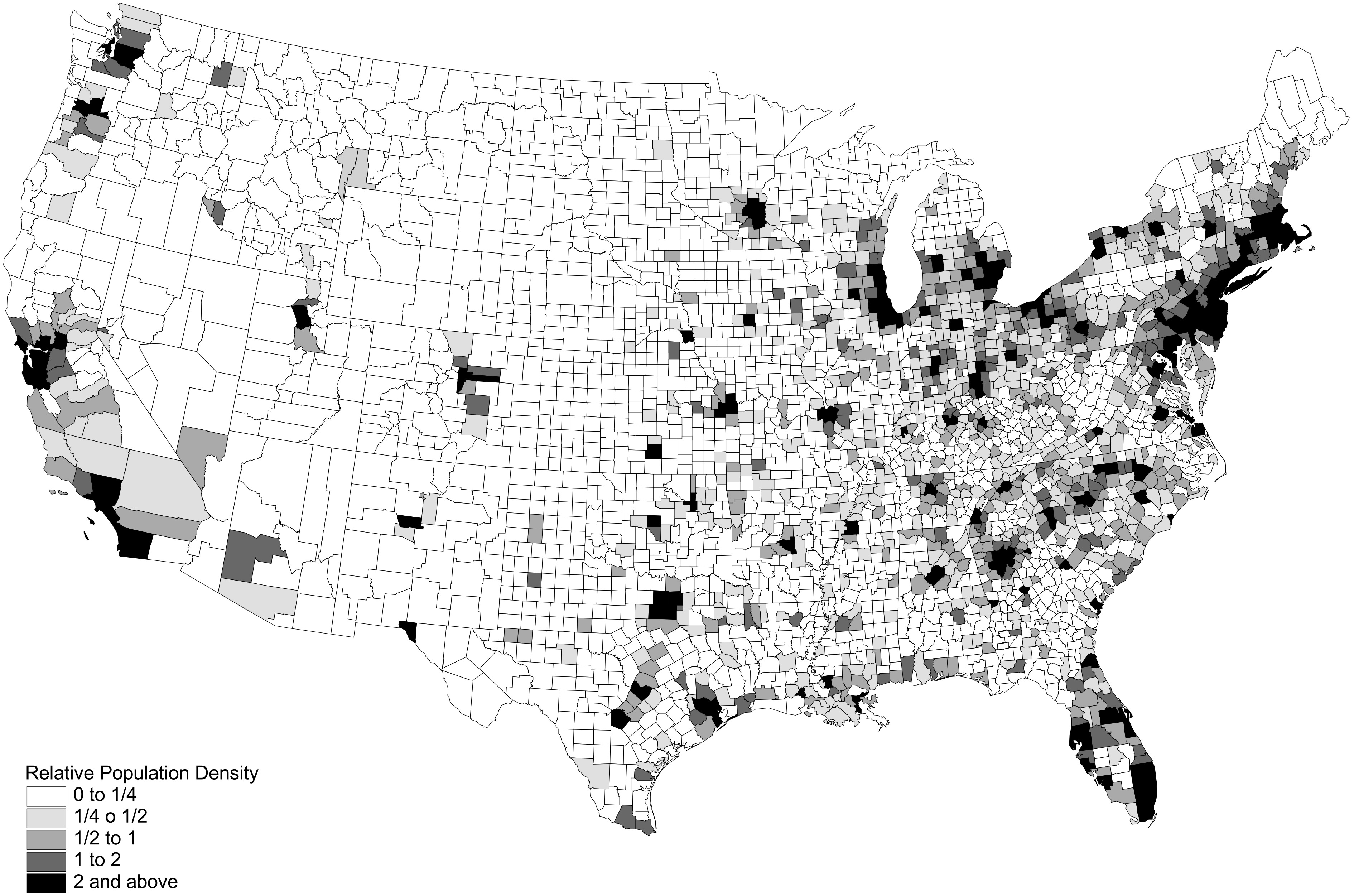
Map 1: Coastal and Navigable River Counties



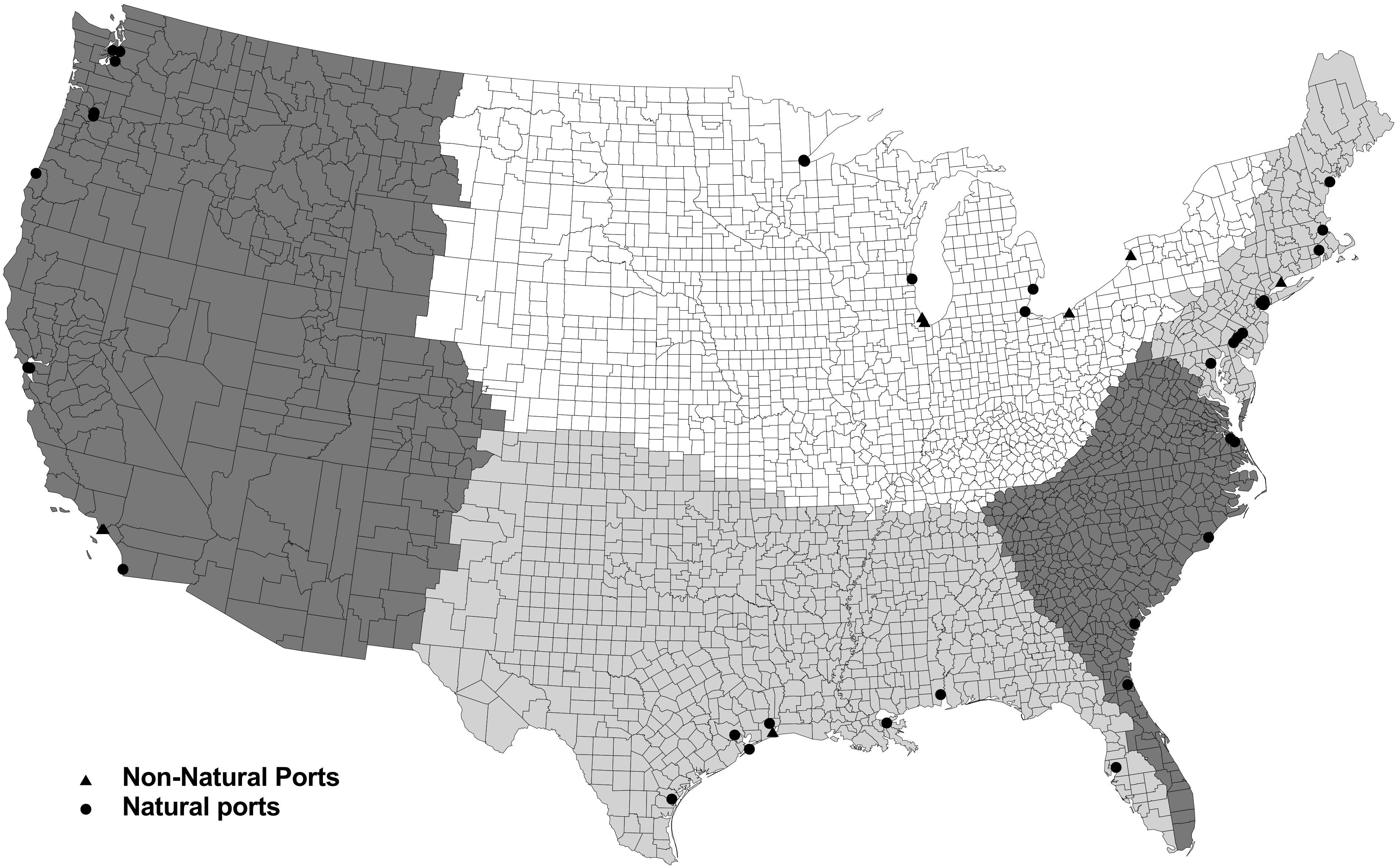
Shaded Counties Represent:
15% of Continental U.S. Land Area
53% of 2000 Population
60% of 1998 Civilian Income

Map 2: Population Density

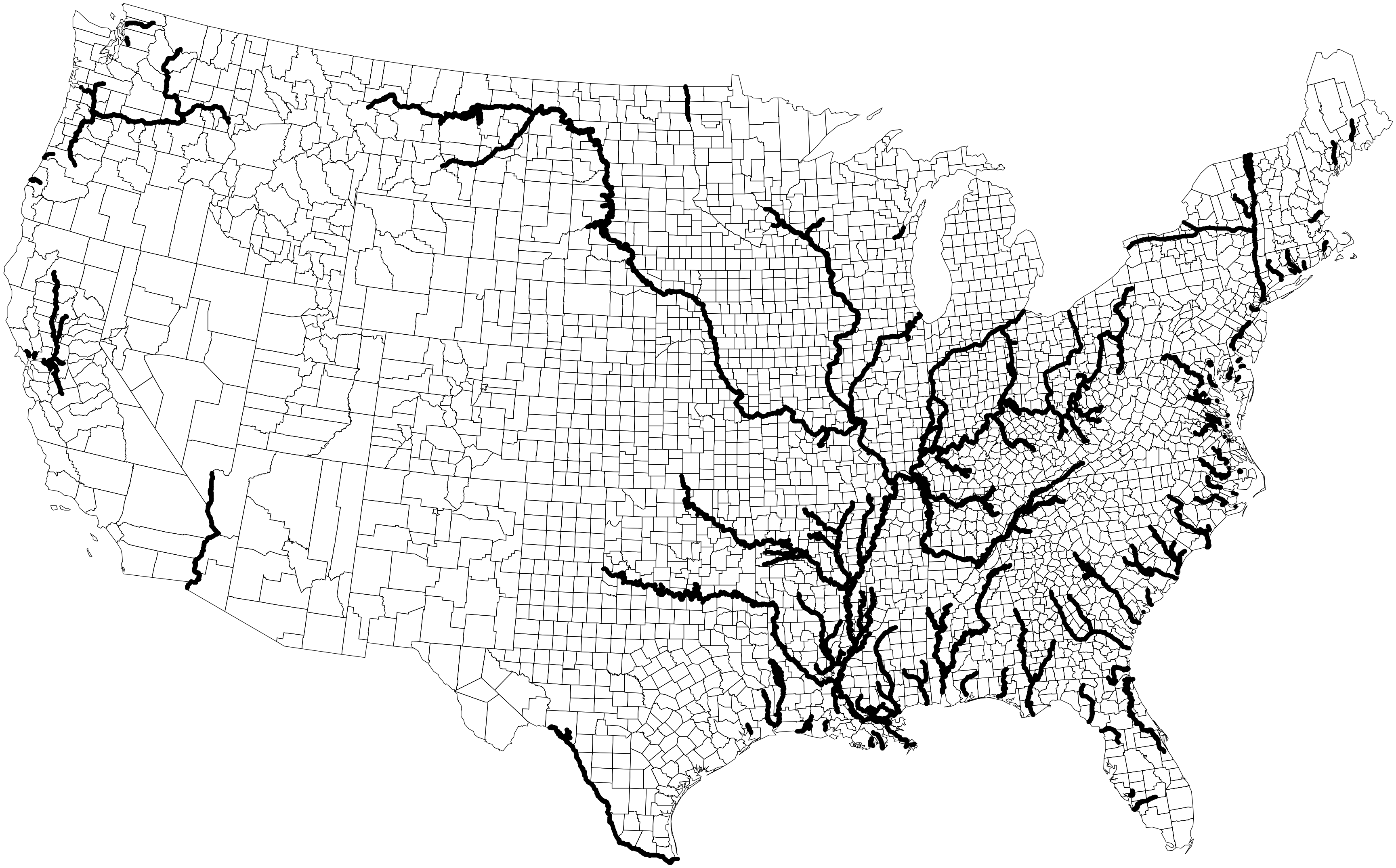
(County Population Density Relative to U.S. Population Density in 2000, 3076 counties)



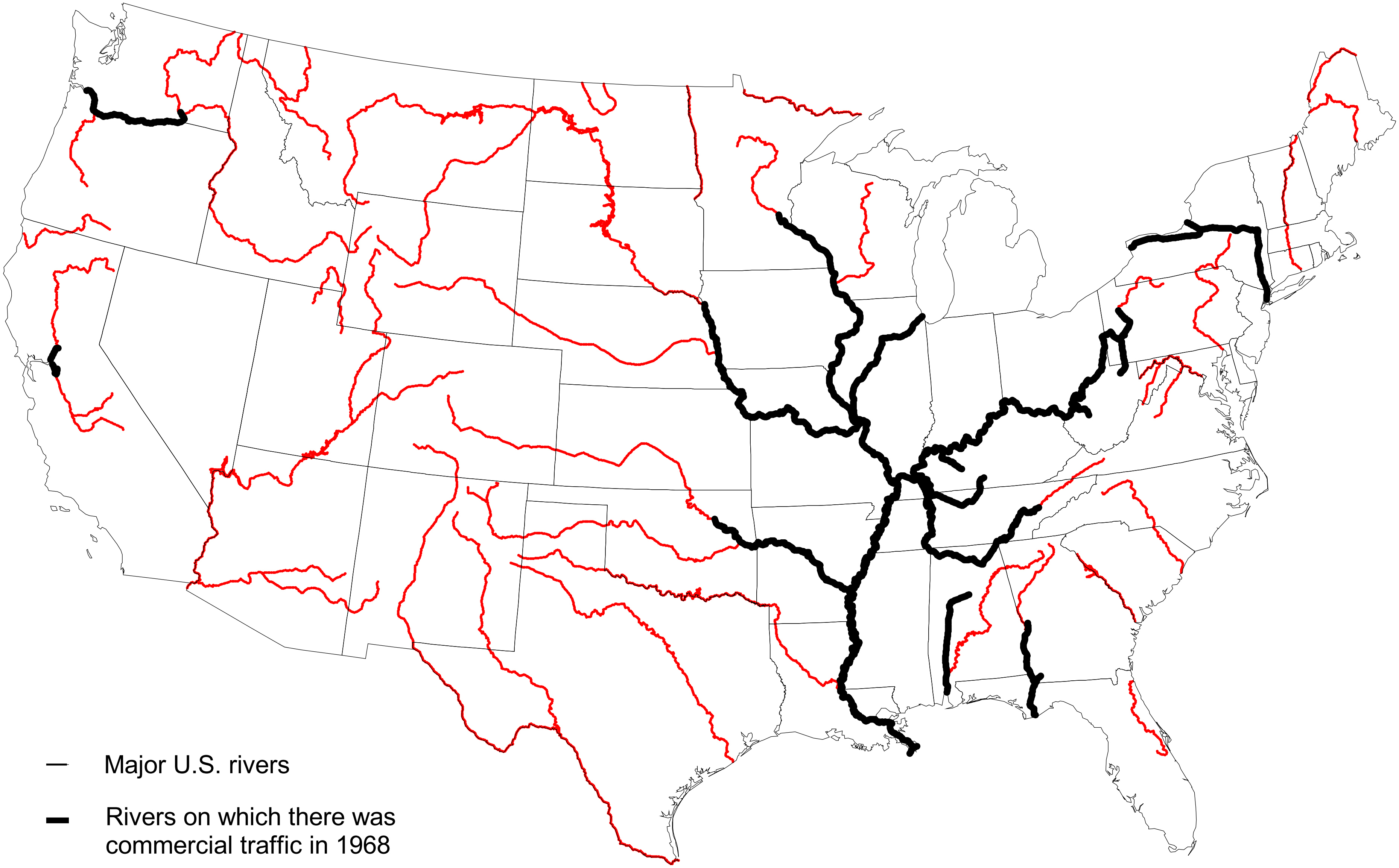
Map 3: Counties and Ports by Nearest Coast



Map 4: Rivers used for navigation in 1890



Map 5 - Navigable and Major U.S. Rivers



— Major U.S. rivers

▬ Rivers on which there was commercial traffic in 1968