Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Regional Science and Urban Economics 37 (2007) 375 – 398
regional SCIENCE & urban ECONOMICS
www.elsevier.com/locate/regec

Moving to nice weather ☆

Jordan Rappaport *

Federal Reserve Bank of Kansas City, 925 Grand Boulevard, Kansas City, MO 64198 USA

Accepted 6 November 2006
Available online 16 January 2007

Abstract

U.S. residents have been moving en masse to places with nice weather. Well known is the migration towards places with warm winters, which is often attributed to the introduction of air conditioning. But people have also been moving to places with cooler, less-humid summers, which is the opposite of what is expected from the introduction of air conditioning. Nor can the movement to nice weather be primarily explained by shifting industrial composition or by migration of the elderly. Instead, a large portion of weather-related movement appears to be driven by an increased valuation of nice weather as a consumption amenity, probably due to broad-based rising per capita income.

JEL classification: R11; R12; R23; N92
Keywords: Economic growth; Population density; Migration; Quality of life

1. Introduction

Over the course of the twentieth century, U.S. residents moved en masse to places with nice weather. Well known is the migration towards places with warm winters, which is often attributed to the introduction of air conditioning (Oi, 1997). But people have also been moving to places with cooler, less-humid summer weather, which is the opposite of what would be expected from the introduction of air conditioning. Nor can the movement to nice weather be attributed primarily to retirement by the elderly, as the trend is nearly as strong among working-age individuals. Instead, regressions of population growth on weather and other characteristics suggest that a large

☆ The views expressed herein are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of Kansas City or the Federal Reserve System. Supplemental materials are available for download from www.kc.frb.org/Econres/staff/jmr.htm.

* Tel.: +1 816 881 2018; fax: +1 816 881 2199.
E-mail address: jordan.rappaport@kc.frb.org.

0166-0462/$ - see front matter
doi:10.1016/j.regsciurbeco.2006.11.004

portion of weather-related moves have followed from a broad-based increase in the valuation of nice weather as a consumption amenity. If so, other consumption amenities are also likely to have served as important sources of local growth. Moreover, if the increased valuations of consumption amenities are the result of rising per capita income, then the migration to nice weather is likely to continue.

Valuing weather's contribution to quality of life has received considerable attention in the compensating differential literature. The value of a weather characteristic can be calculated as the sum of the wages a household is willing to forego plus the price premium they are willing to pay to live in a place with that characteristic (Rosen, 1979; Roback, 1982). Based on public-use micro-samples from the 1980 decennial census, one hundred fewer annual heating degree days are estimated to be valued (in 2002 dollars) from \$5 to \$40 per household; one hundred fewer annual cooling degree days are estimated to be valued from \$2 to \$218 per household; one extra sunny day per year, from \$19 to \$33 per household; and one inch less precipitation, from −\$58 to \$34 (Blomquist et al., 1988; Gyourko and Tracy, 1991; Stover and Leven, 1992). More recent research finds that Italian households are similarly willing to pay substantial sums to enjoy less precipitation and more moderate summer temperatures (Maddison and Bigano, 2003). Other recent research finds a substantial increase during the latter part of the twentieth century in U.S. households' valuation of weather's contribution to quality of life (Cragg and Kahn, 1999; Costa and Kahn, 2003).

There are several important limitations to the compensating differential methodology of valuing weather. First is that the number of U.S. geographic observations is relatively small, since the necessary micro data is available only for places with a population of at least 100,000. These observations are subject to a sharp selection bias since places where large numbers of people have not chosen to live are excluded. A second limitation of the compensating differential methodology is the difficulty controlling for individual-specific and house-specific characteristics (Gyourko et al., 1999; Combes et al., 2004; Lee, 2005). To the extent that higher-income individuals disproportionately desire to live in high quality-of-life locations, there will be a strong positive correlation between local consumption amenities and respondents' unobserved characteristics. A third limitation is the identifying assumption that the system of localities across which attributes are being valued is at its long-run steady state. Highly persistent population flows among localities suggest a system in continuous transition (Greenwood et al., 1991; Rappaport, 2004a).

In contrast to the "price" approach of the compensating differential literature, the present paper pursues a "quantity" approach. Partial correlations of population growth with local weather reflect changes in the valuation of weather's combined contributions to quality of life and productivity. Continental U.S. counties are used as the geographic unit of observation. The large number of these – more than three thousand – allows for high statistical power. The discrete partition of the continental United States captures an extremely wide variation in weather. And the underlying theory explicitly takes into account that steady-state and actual population are likely to differ.

There are three main limitations to the quantity approach. First is that it can generally capture only *changes* in contributions to quality of life and productivity. Second is that it does not quantify the size of the changed contributions. Third is that it fails to distinguish between changed contributions to quality of life versus changed contributions to productivity.

The paper proceeds as follows: Sections 2 and 3 respectively discuss the paper's theoretical and econometric framework. Section 4 briefly describes the data. Section 5 reports and interprets the empirical results. A last section concludes.

2. Theory

Two sets of theory underpin the paper's empirics. The first concerns how to interpret partial correlations between local population growth and exogenous attributes. The second considers specific ways in which the contribution of weather to quality of life and productivity has changed over the past century.

2.1. The determinants of local population

The theoretical framework for inferring the determinants of local population follows Haurin (1980) and Rappaport (2004a,b). The world is assumed to comprise a large number of "localities." These are fixed geographic areas where people both live and work and across which there is at least moderate factor mobility. Localities *exogenously* differ with respect to some attributes that affect local quality of life and local productivity. Local attributes that affect quality of life enter directly as arguments in individuals' utility functions; those that affect productivity enter directly as arguments in firms' profit functions. Equivalently, quality of life attributes shift individuals' indifference curves over wages and prices. Productivity attributes shift firms' iso-profit curves over wages and other input prices.

In a long-run steady state, population density is an increasing function of local quality of life and local productivity. Individuals are willing to endure greater crowdedness and the associated higher price of housing in order to directly enjoy higher quality of life and to indirectly enjoy higher productivity via the higher wages it affords. Indeed, varying local population density is the primary mechanism by which local wages and house prices adjust to equate utility and profits across localities.

The setup might suggest that a "level" regression of population density on exogenous attributes can identify their contributions to quality of life and productivity. But this is unlikely: factor mobility is not perfect, and both productivity and quality of life also depend on endogenous attributes.

Imperfect factor mobility implies that current population density may differ substantially from steady-state population density. Even relatively small frictions to labor and capital mobility can cause population to require several decades to transition from one steady state to another (Rappaport, 2004a). Hence a positive partial correlation between current population density and an exogenous attribute may reflect a past rather than a present contribution to quality of life and productivity.

The partial dependence of quality of life and productivity on endogenous attributes introduces an additional difficulty: such endogenous attributes are likely to be correlated with the exogenous ones. Substantial research suggests that quality of life and productivity themselves depend on population size and density. Size and density may increase quality of life by allowing for increased social interaction and product variety (e.g., Glaeser et al., 2001; Fujita and Thisse, 2002; Compton and Pollak, 2004). And they may increase productivity via local scale economies (e.g., Henderson et al., 1995; Ciccone and Hall, 1996; Henderson, 2003). An exogenous attribute that increased quality of life and productivity in the past may no longer do so in the present. But the high population density caused by the past contribution may cause *steady-state* population density to remain correlated with the attribute via a "lock-in effect" from self-reinforcing agglomeration forces (Fujita and Mori, 1996).

Interpreting partial correlations between population growth and exogenous attributes is more straightforward. Regardless of factor mobility, an exogenous attribute that is partially correlated with long-run population growth will be partially correlated with a change in steady-state density. Imperfect factor mobility simply implies that the change may have occurred substantially before the period over which population growth is being measured. One possible cause of a change in steady-state density is a change in the contribution from the exogenous attribute to quality of life and productivity. Another is a change in the contribution to these from an endogenous attribute that is correlated with the exogenous attribute. To rule out the latter possibility, regressions can include controls for predetermined endogenous attributes.

2.2. The changing contributions from weather

That the weather contributes to quality of life and productivity is obvious. Indeed, it is hard to think of any outdoor activity – whether for leisure or for work – that is not affected by weather. But for present purposes, the relevant question is how weather's contribution to quality of life and productivity has changed over the past century. Among many possible ways, four stand out. The first is that the decline of agricultural employment has decreased the importance of weather characteristics that contribute to agricultural productivity. The second is that air conditioning and improved heating technology have decreased the disutility from extreme temperatures. The third is that rising per capita income has increased the relative valuation of temperate weather. The fourth is that it is primarily the growing number of affluent and mobile retirees who have increased their valuation of temperate weather.

The explanation based on declining agricultural employment is straightforward. Agricultural productivity clearly depends on numerous weather attributes, such as rainfall and the length of a growing season. While this dependence may have remained largely unchanged over the course of the twentieth century, the importance of agricultural productivity to overall U.S. productivity has become much smaller. Agriculture's share of U.S. employment fell from approximately 40% in 1900 to less than 2% in 2000. The sector's relative decline freed up more than a third of U.S. residents and jobs to relocate based on considerations other than agricultural productivity. A related possibility is that relative employment declines in the minerals extraction and manufacturing industries freed up people to move away from bad-weather places where such industries were concentrated.

The roles of air conditioning and improved heating technology require a bit more explanation. As a matter of background, the first known air-conditioning units were installed around 1900. Over the subsequent 40 years, AC was slowly adopted by manufacturers and a few service businesses. However, it was not until after World War II that the mass adoption of residential air conditioning began. As late as 1960, only 13% of U.S. households had any sort of AC and only 2% had central AC. Home heating also improved dramatically over the course of the twentieth century. Starting in the 1920s, vendor-delivered oil came to supplant coal as the primary fuel source. Starting in the 1940s, utility-supplied natural gas and electricity came to supplant oil. During the latter part of the century, there were also considerable improvements in home insulation technologies.[1]

To help understand the implications of air conditioning and improved heating, Fig. 1 shows average daily high summer and winter temperatures for 45 U.S. urbanized areas. The summer measure is a discomfort index reflecting both temperature and humidity. More details on the weather variables are in the data-description section below. The solid lines in the figure represent a possible set of indifference curves prior to AC and improved heating. Utility is decreasing as

[1] Tables documenting the incidence of air conditioning and primary heating fuel across states and decades are included in the paper's supplemental materials, which are available for download from www.kc.frb.org/Econres/staff/jmr.htm.

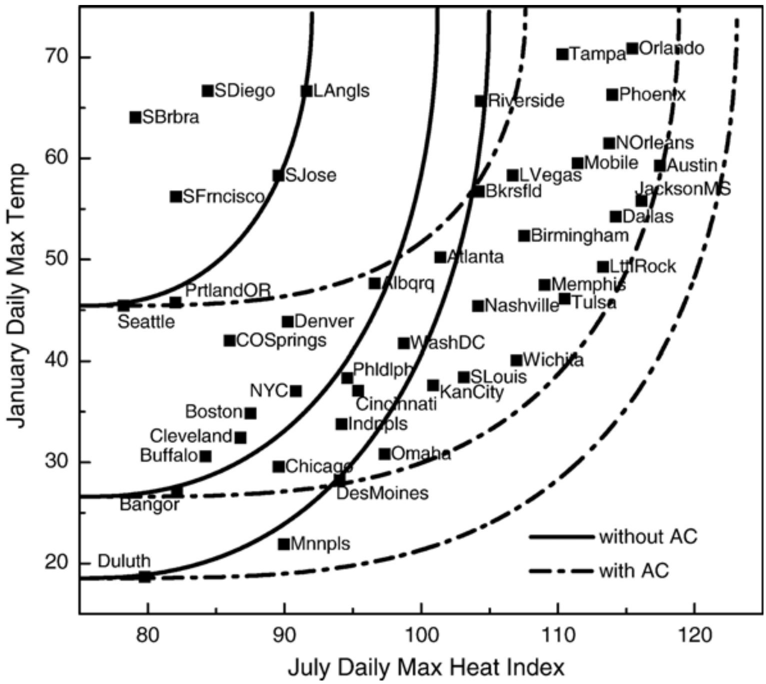

Fig. 1. Iso-utility over weather.

you move to the lower right. The underlying loss function assumes 75 °F to be the year-round ideal.[2]

The introduction of air conditioning lessens the disutility from hot summer weather. Individuals become willing to endure more of it in return for warm winter weather. Equivalently, the indifference curves stretch rightward along the top vertical axis. This causes numerous preference reversals. For example, Duluth Minnesota's weather without air conditioning is preferred to that of numerous cities in the South and Southwest. But with AC, Duluth's weather is the least preferred among the depicted cities. Similarly, numerous cities in the bottom left of the figure reverse from having relatively desirable to relatively undesirable weather.

Restoring spatial equilibrium requires a new vector of wages and house prices across places. In general, the more the valuation of a place's weather has risen, the larger the required fall in its relative wage and increase in its relative house price. Migration brings about the new spatial equilibrium. People will move both towards places with hot summers and places with warm winters (that is, towards the upper right of the figure). The impulse towards hot summers is intuitive: hot weather has become "less bad". The impulse towards warm winters arises because AC increases the sensitivity of relative weather valuations to winter temperatures. This is reflected in the flatter indifference curves.[3]

[2] Specifically, the loss function is assumed to be of the form, $L=a|s-s^*|^{\alpha}+b|w-w^*|^{\beta}$. Here s and w are respectively summer and winter temperature. The remaining elements of the equation are positive parameters. The introduction of AC is depicted as a decrease in a. Improved heating might correspond to a decrease in b. An increased valuation of nice weather might correspond to *increases* in both a and b.

[3] An illustrative example, in the upper middle of Fig. 1, is the required faster relative growth of Riverside compared to Bakersfield, notwithstanding that the two cities' summer weather is virtually identical. Prior to AC, the two cities' weather bundles were equally valued. With AC, Riverside's weather is preferred to that of Bakersfield.

Reinforcing the faster relative growth of cities with warm winters is the strong positive correlation between summer and winter temperatures.

Improved heating technologies cause indifference curves to shift in a way that mirrors the shift from air conditioning. Specifically, the curves stretch downward along the left vertical axis. As a result, people move both towards places with cold winters and towards places with cool summers (that is, towards the lower left of the figure).

Although air conditioning and improved heating imply migration in the opposite direction, the two technological changes in combination only partially offset each other. With approximately equal magnitude shifts from air conditioning and from improved heating, the AC-induced impulse toward warm winters and the heating-induced impulse towards cool summers are indeed cancelled out. But there continues to be a combined impulse toward cold winters and hot summers. To see this, suppose that the combined technological shocks cause indifference curves to shift downward and rightward such that the new curves exactly parallel the old ones. As a result of such a shift, the required compensating variation for any given deviation from ideal weather falls. Restoring spatial equilibrium thus requires people to move in the same direction as the shift in indifference curves.

In sharp contrast to the implied movement from air conditioning and improved heating, the rise in income associated with broad-based technological progress creates an impulse for people to move toward places with nicer weather. Over the course of the twentieth century, rising total factor productivity and the capital accumulation it implied increased U.S. per capita income more than six-fold (Maddison, 1995). To the extent that nice weather is a normal good, demand for it should have risen as well. More specifically, increased consumption of goods is associated with a decrease in marginal utility. Hence, people should be willing to pay a higher price, in terms of foregone consumption, to live in a place with nice weather. In terms of Fig. 1, rising incomes cause indifference curves to shift inward: a given loss occurs at a smaller deviation from ideal weather. Maintaining spatial equilibrium thus requires places with nice weather to have an increasingly negative wage premium and an increasingly positive house price premium. Migration to places with nice weather can help bring about the new spatial equilibrium by putting downward pressure on wages and upward pressure on house prices.

However, broad-based technological progress can also induce substitution effects that offset the income effect just described. Suppose, for example, that technological progress causes an equal-proportional rise in wages across all localities. In absolute terms, the wage penalty to living in a high quality-of-life locality will have increased. Similarly, technological progress may be relatively slower for the production of housing than for other goods. As a result, suppose that house prices rise by an equal proportion across all localities. In absolute terms, the housing price penalty to living in a high quality-of-life locality will have increased. In a general equilibrium model with CES production and utility, the direction of movement from rising wages depends on nontrivial parameter restrictions. As a rule of thumb, movement will be toward nice weather if the elasticity of substitution among the sources of utility is below one (Rappaport, 2004b). The intuition is similar to that of the exact balancing of income and substitution effects with log utility such that rising wages do not increase leisure consumption.

The final hypothesized way in which the contribution from weather to quality of life and productivity has changed is similar to rising per capita income, except that it applies only to the elderly. Social Security and better retirement planning have increased the wealth of retirees. Medical advances have increased their longevity. And many affluent individuals are deciding to retire early. Because they do not work, retirees' income no longer depends on their locale's productivity. And so upon retiring, quality of life suddenly becomes a much more important determinant of where they live.

The empirical section below argues that growth regressions can distinguish, at least in part, among these four possible changes in weather's contribution to quality of life and productivity. Controlling for predetermined agricultural, mineral extraction, and manufacturing characteristics of counties suggests that the movement towards nice weather does not arise primarily from a shift in industrial composition. The faster relative growth of places with cooler, less humid summers establishes that air conditioning cannot alone account for the movement toward nice weather. And the approximately equal draw of nice weather for both working-age and elderly individuals suggests that increased retirement has been a relatively unimportant source of the population shift. Overall, the empirics are most consistent with the explanation that individuals have been steadily increasing their valuation of nice weather's contribution to quality of life. In addition, fitted growth rates suggest that air conditioning has also played a large role in the population movement.

3. Econometric framework

A locality's steady-state population density is assumed to depend on its productivity and quality of life. In turn, these are assumed to depend on actual density, on time-invariant attributes, and on time-varying local attributes. More specifically, the natural log of steady-state population density for locality i is assumed to take the linear functional form,

$$l^*_{i,t} = \alpha_t l_t + \mathbf{x}_i^{'}\boldsymbol{\beta}_t + \mathbf{y}_{i,t}^{'}\gamma_t + \mu_{i,t} \quad 0 \leq \alpha < 1 \tag{1}$$

The term α_t is a time-varying scalar that captures any increasing returns from current population density to quality of life and productivity. Its restriction to be less than one assures a stable system of cities rather than the concentration of all population in a single city. The k-by-1 column vector $\mathbf{x}_i$ measures the locality's time-invariant attributes. These are multiplied by a time-varying k-by-1 column vector, $\boldsymbol{\beta}_t$. The j-by-1 column vector $\mathbf{y}_{i,t}$ measures local time-varying attributes that affect quality of life and productivity. These are multiplied by a time-varying j-by-1 column vector, $\boldsymbol{\gamma}_t$. The stochastic disturbance, $\mu_{i,t}$, is assumed to have expectation zero and to be uncorrelated with the remaining right-hand-side variables.[4]

Only actual density rather than steady-state density is observed:

$$l_{i,t} = \alpha_t l_t + \mathbf{x}_i^{'}\boldsymbol{\beta}_t + \mathbf{y}_{i,t}^{'}\gamma_t + \mu_{i,t} + (l_{i,t} - l^*_{i,t}) \tag{2}$$

An expression for population growth follows from first-differencing (2).[5] An error term, ε, includes components that may be correlated with $\mathbf{x}_i$.

$$l_{i,t} - l_{i,t-1} = \left(\frac{1}{1-\alpha_t}\right)((\alpha_t - \alpha_{t-1})l_{i,t-1} + \mathbf{x}_i^{'}(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}) + \mathbf{y}_{i,t-1}^{'}(\gamma_t - \gamma_{t-1})) + \varepsilon_{i,t} \tag{3}$$

$$\varepsilon_{i,t} = \left(\frac{1}{1-\alpha_t}\right)((\mathbf{y}_{i,t}^{'} - \mathbf{y}_{i,t-1})\gamma_{t-1} + (\mu_{i,t} - \mu_{i,t-1}) + \underbrace{((l_{i,t} - l^*_{i,t}) - (l_{i,t-1} - l^*_{i,t-1})))}_{\textit{convergence component}} \tag{4}$$

[4] The label "steady state" might alternatively be reserved for a stronger condition characterized by substituting $l^{**}_{i,t}$ for both $l^*_{i,t}$ and $l_{i,t}$ in (1). Separately, x_i is assumed to include a 1 as its first element, thereby allowing for a time-varying constant.

[5] The growth rates of total population and population density are equal so long as a place's land area remains constant. The two growth rates are referred to interchangeably herein.

The growth regressions that follow have reduced functional form,

$$l_{i,t} - l_{i,t-1} = \alpha l_{i,t-1} + \mathbf{x}_i' \mathrm{b} + \mathbf{y}_{i,t-1}' \mathrm{g} + \varepsilon_{i,t} \tag{5}$$

For present purposes, we are most interested in interpretations of the coefficients, b, on time-invariant variables. To motivate such interpretations, it is helpful to temporarily assume that the convergence component of ε always equals zero. This would be true either if actual density always equaled steady-state density or if the difference between actual and steady-state density remained constant over time.

The simplest growth regressions below include only time-invariant attributes, $\mathbf{x}_i$. In addition to zero convergence, assume that there are no increasing returns to scale (α_t always equal to zero) and that any excluded time-varying variables that affect steady-state density are orthogonal to the time-invariant ones. In this case, b will give an unbiased estimate of $(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})$.

Reintroducing increasing returns to scale broadens interpretations of b. As a first step, suppose that α_t is positive but constant over time, $0 < \alpha < 1$. In this case, b is an unbiased estimate of $1/(1-\alpha)$ times $(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})$. Increasing returns "magnifies" the effect of any change in contribution from the time-invariant attributes to productivity and quality of life.

Now allow the degree of increasing returns to vary over time. So, for example, an increase in α_t will increase the steady-state density of places that already have high density. In this case, including initial density on the right hand side avoids an omitted variable bias. Otherwise, it would falsely appear that time-invariant attributes' contributions to quality of life and productivity had increased, when in fact they had remained constant. That is, b would estimate $\rho \boldsymbol{\beta}_{t-1}$ for some ρ greater than zero. A decrease in α would cause b to estimate the same for some ρ less than zero.

Similar to the case of initial population density, there may be other time-varying attributes, $\mathbf{y}_{i,t}$, that are correlated with $\mathbf{x}_i$ and whose contributions to quality of life and productivity have changed over time. A possible example is the local stock of manufacturing capital. It is likely to be correlated with weather due to the importance of proximity to iron ore, coal, and other raw materials concentrated in the U.S. Midwest. The ongoing decline of manufacturing's share of U.S. employment can be interpreted as a decrease in the productivity contribution from manufacturing capital. Including the *initial level* of manufacturing capital then ameliorates an important omitted-variable bias. Even more ideal would be to explicitly control for time-invariant attributes, such as access to raw materials, for which manufacturing capital may proxy.

On the other hand, the inclusion of time-varying attributes in Eq. (5) may mask changes in contributions by time-invariant ones. Any particular element $y_{i,t-1}^k$ may be an endogenous response to more fundamental underlying quality-of-life and productivity differences. The sharp decline in agricultural employment over the twentieth century can be interpreted as a decline in the contribution to overall productivity from time-invariant characteristics – including the weather – that increase agricultural productivity. Since agricultural intensity should be highly correlated with such characteristics, controlling for its initial level may cause a negative estimate of some g^j in place of a negative estimate of some corresponding b^k.

Attempting to control for *changes* in time-varying attributes would make this problem of endogeneity even worse. In addition to depending on time-invariant attributes as just described, time-varying attributes may also depend on population density itself. Because Eq. (5) already controls for initial density, this additional source of endogeneity from the inclusion of $\mathbf{y}_{i,t-1}'$ is unlikely to seriously bias estimates of $(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1})$. In contrast, $(\mathbf{y}_{i,t} - \mathbf{y}_{i,t-1})$ may absorb a large share of the variation that is actually caused by exogenous right-hand-side variables. For this reason, the causal contribution from changes in time-varying attributes to population growth is relegated to the error term, $\varepsilon_{i,t}$. Interpretations of estimated coefficients must consider the possible bias from this omission.

A final complication to interpreting coefficients comes from relaxing the assumption that the convergence component in Eq. (4) remains constant. In the absence of perfect labor mobility, this assumption will be false. Moreover, the change in population density's distance from its steady state is almost certainly correlated with the included right-hand-side variables.

The most obvious bias comes from the slow adjustment of population density following a contemporary change in structural coefficients. Consider a positive change in the structural coefficient on some particular time-invariant attribute, $(\beta_t^k - \beta_{t-1}^k) > 0$. Since population may take several decades to transition to its new steady state, the regression coefficient on x^k underestimates the change in structural coefficients. In this case, the convergence term, $((l_{i,t} - l_{i,t}^*) - (l_{i,t-1} - l_{i,t-1}^*))$, is negatively correlated with x_i^k. Measuring growth over a long interval does not necessarily eliminate this bias. In the main set of regressions below, growth is measured from 1970 to 2000. But even if thirty years is sufficient for density to converge to a new steady state, nevertheless discrete changes in structural coefficients may have occurred anytime over the interval, including towards the very end.

A bias in the opposite direction occurs following a change in structural coefficients that precedes the interval over which growth is measured. Suppose that $(\beta_t^k - \beta_{t-1}^k) = 0$ but that $(\beta_{t-1}^k - \beta_{t-2}^k) > 0$. During the interval over which growth is measured, density is still adjusting to its new steady state. Hence the estimated b^k is positive. In this case, the convergence term, $((l_{i,t} - l_{i,t}^*) - (l_{i,t-1} - l_{i,t-1}^*))$, is positively correlated with x_i^k.

Together, these two biases arising from slow convergence caution against interpreting coefficients in the growth regressions as literally estimating the *magnitude* of changes in structural coefficients. Instead, b should be interpreted as estimating an accumulation of past changes, $\sum_{\tau=0}^{T} \delta_\tau (\boldsymbol{\beta}_{t-\tau} - \boldsymbol{\beta}_{t-\tau-1})$ with unknown weights, δ_τ, and unknown starting period, $(t-T-1)$. In particular, the sign of any specific estimated b^k can be interpreted as indicating the *direction* of recent changes in β_t^k.

Regressions using county observations almost surely violate the classical assumption of independence. This is due to possible spatial components both of omitted variables and of stochastic disturbances. Hence a generalization of the Huber–White heteroskedastic-consistent estimator is used to report standard errors that are robust to a spatial structure among regression residuals (Conley, 1999). Let $\sigma_{i,j}$ be an element of the covariance matrix $E(\boldsymbol{\varepsilon\varepsilon}')$. For pairs of observations between which the Euclidean distance is beyond a certain cutoff, covariance between residuals is assumed to be zero. For pairs of observations closer to each other than the cutoff, a declining weighting function is imposed to estimate residual covariance. Let $s_{i,j}$ be the estimate of $\sigma_{i,j}$, and let ε_i be a regression residual.

$$s_{i,j} = \mathrm{g}(\mathrm{distance}_{i,j})\varepsilon_i\varepsilon_j \tag{6}$$

$$\begin{aligned} &\mathrm{g}(\mathrm{distance}_{i,j}) \begin{cases} = 1 & \mathrm{distance}_{i,j} = 0 \\ \in [0,1] & 0 < \mathrm{distance}_{i,j} \le \overline{d} \\ = 0 & \mathrm{distance}_{i,j} > \overline{d} \end{cases} \\ &\mathrm{g}'(\mathrm{distance}_{i,j}) \quad \le 0 \end{aligned} \tag{7}$$

Herein, the weighting on the estimated covariance between residual terms is assumed to fall off quadratically as the distance between county centers increases to 200 km.[6] Thus accounting for spatial correlation approximately doubles standard errors relative to the assumption of zero covariance with homoskedastic disturbances.

4. Data description

U.S. counties are used as the geographic unit of analysis. Doing so offers several benefits relative to using alternative U.S. local geographies. First, counties completely partition the continental United States. Excluding geographic areas with low population would introduce a source of considerable bias. Second, counties' borders have been relatively constant over time. Constant borders allow intertemporal comparisons between geographically fixed areas that can be considered exogenous relative to most data-generating processes.[7]

The primary dependent variable below is the annual growth rate of population density from 1970 to 2000. Other regressions use the growth rates of population subgroups and house prices. Summary statistics for the growth rates are included in Table 1. The variables are constructed based on U.S. Census Bureau (1890-1960, 1947-1977, 1980, 1990, 2000a,b) and U.S. Bureau for Economic Analysis (1969-2000a,b).

The weather variables are derived from data purchased from The Climate Source, Inc. (www.climatesource.com). The Climate Source data, in turn, is based on detailed weather observations over the period 1961 to 1990 from more than 6000 meteorological stations managed by the U.S. National Oceanographic and Atmospheric Administration. A peer-reviewed "hybrid statistical–geographical methodology" is applied to such data to fit surfaces over a two-kilometer grid of the continental United States. A county's weather values are then constructed as the mean over all of the four-square-kilometer grid cells that lie within it. In the regressions, weather variables enter with both a linear and quadratic term. The quadratic terms have had their linear sample mean subtracted prior to squaring. Doing so allows the coefficient on the linear term to measure the marginal effect of an increase in the variable from the sample mean.

Winter weather is measured by January daily maximum temperature. This is the average maximum temperature for days in January; in other words, it is the average of the warmest temperature attained on each of the 930 January days from 1961 to 1990. The choice to use January daily maximum rather than minimum temperature reflects an a priori belief that winter daytime highs are likely to be a more important contributor to quality of life than winter nighttime lows. Results are extremely robust to using alternative winter temperature measures.

Summer weather is measured by July daily maximum heat index and July average daily mean relative humidity. The latter is the mean of average daily maximum and minimum humidity. The former is a discomfort index that combines average daily maximum temperature with average daily mean relative humidity (Stull, 2000). The inclusion of humidity entered independently in addition to its contribution to heat index is motivated by its strong marginal power to account for population growth's sample variance. Results are extremely robust to using alternative summer temperature measures.

[6] In other words, $g(\cdot) = 1 - \left(\frac{\text{distance}_{i,j}}{200}\right)^2$. Note that the present specification reduces to the Huber–White heteroskedastic-consistent estimator for standard errors when $\bar{d}$ equals zero; it reduces to a group-based random-effect estimator for standard errors with a non-Euclidean one–zero step specification for g(·).

[7] Nevertheless, a few adjustments to county geographies are required. Details are described in Rappaport and Sachs (2003) based on Horan and Hargis (1995).

Table 1
Summary statistics

Variable	Obs	Mean	Std. dev.	Min	Max
1970-to-2000 growth rates					
Population	3067	0.9	1.3	−2.4	10.1
Employment (civilian)	3065	1.8	1.5	−2.8	11.3
Working age pop. (aged 25 to 54)	3067	1.7	1.4	−2.4	11.3
Elderly population (aged 65 and up)	3066	1.7	1.4	−2.2	10.0
Native U.S. citizen	3067	0.8	1.3	−2.5	10.0
Immigrant	3067	3.5	4.3	−8.3	24.4
Median house value	3063	2.3	0.9	−3.7	7.7
Median house rent	3049	1.4	0.7	−1.2	4.2
Population growth rates by decade					
1880s	2395	3.5	7.0	−13.2	69.2
1890s	2604	1.9	3.2	−11.5	37.9
1900s	2696	2.0	4.1	−8.8	45.6
1910s	2844	0.7	2.2	−21.9	30.9
1920s	3014	0.7	2.8	−8.9	44.3
1930s	3060	0.5	1.5	−6.4	15.0
1940s	3062	0.3	2.0	−9.3	14.5
1950s	3064	0.4	2.2	−5.5	15.7
1960s	3063	0.4	1.7	−11.3	12.1
1970s	3067	1.4	1.7	−4.5	13.0
1980s	3067	0.3	1.4	−3.9	9.7
1990s	3069	1.0	1.3	−4.7	10.7
Weather					
January daily max. temp. (°F)	3067	41.4	12.4	10.8	76.5
July daily heat index (°F)	3067	98.3	11.1	75.6	131.3
July daily relative humidity	3067	65.7	9.6	23.8	82.0
Annual precipitation	3067	38.3	14.1	3.5	118.2
Annual precipitation days	3067	94.1	24.5	13.3	198.0

All growth rates are on an annual percentage basis.

All regressions also include controls for precipitation. These are made up of linear and quadratic terms for average annual precipitation and for the average number of days per year on which there was at least 0.01 inch of precipitation. Maps depicting the distribution of each of the five weather elements across U.S. counties are included in the paper's supplemental materials.

Some of the regressions include a set of seven geographic controls measuring coastal proximity and topography. Separate dummies indicate whether a county's center is within 80 km of an ocean or Great Lakes coast or within 40 km of a major river. Additional variables measure ocean and Great Lakes shoreline per unit area. And a topography variable, entered linearly and quadratically, measures the standard deviation of altitude across 1.25-arc-minute grid cells within a county divided by total county land area.

To control for increasing returns to scale, some of the growth regressions include an additional set of fourteen variables measuring initial population density and surrounding total population. Initial population density is entered as a seven-part spline to allow for a nonlinear relationship between initial and steady-state population density. Surrounding total population is the initial total population in seven concentric rings emanating from a county's center. An innermost circle measures (the natural log of) total population of all counties with centers within 50 km of a county's own center. At a minimum, this innermost circle always includes the county's own population. A second ring measures the total population of all counties with centers 50 to 100 km from a county's own center. Additional rings have outer-circle radii of 150 km, 200 km, 300 km, 400 km, and 500 km. Together, these concentric population variables capture, for instance, the "market potential" available to local firms that produce goods with nontrivial transportation costs (Krugman, 1991; Ades and Glaeser, 1999; Fujita et al., 1999; Hanson, 2001; Black and Henderson, 2002).

As discussed in the Theory section, one of the explanations for the movement to nice weather is that it reflects the decline of agricultural, mineral extraction, and manufacturing employment. A set of seventeen variables, listed in Table 2, controls for this. The industry-share variables are constructed from the 1970 decennial census. The remaining variables are constructed from data collected for the 1969 and 1972 economic censuses and disseminated in the U.S. Census Bureau Consolidated City and County Databook (ICPSR Study 7736).

5. Results

Throughout much of the 20th century, local population growth in the United States has been positively partially correlated with winter temperature and negatively partially correlated with summer temperature and summer humidity. Supporting results are presented and interpreted for three sets of regressions. The first uses 1970-to-2000 population growth as its dependent variable and incrementally introduces different categories of controls. The second set applies a base specification from the first set to several alternative dependent variable growth rates. The third set applies a base specification to decade growth rates from the 1880s through the 1990s. Together the results suggest that rising incomes are likely to have played a large role in causing the move to nice weather. They also suggest that air conditioning was probably important as well.

5.1. Population growth, 1970 to 2000

Results of the 1970-to-2000 growth regressions are reported in Table 3. In Column 1, the weather variables are included without any additional controls. A positive coefficient on January temperature statistically differs from zero at the 0.05 level. It implies that expected population growth increases as average daily maximum temperature increases above its sample mean value of 41 °F. The positive, statistically significant coefficient on quadratic January temperature implies that the higher expected population growth from warmer winter temperatures becomes

Table 2
Agriculture, manufacturing, and mineral extraction control variables

Industry share of 1970 total employment (10)	Additional agriculture controls (3)
	Farming occupation share (1969)
Agriculture	Percent of land devoted to farming (1969)
Mining	Agriculture sales per capita (1969)
Primary metals manufacturing	
Fabricated metals manufacturing	Additional manufacturing controls (2)
Machinery manufacturing	Manufact. industries payroll per capita (1972)
Transportation equip. manufacturing	Manufact. industries val. added per capita (1972)
Other durable manufacturing	
Textiles manufacturing	Additional mineral controls (2)
Chemicals manufacturing	Mineral industries payroll per capita (1972)
Other nondurable manufacturing	Mineral industries sales per capita (1972)

Table 3
Population growth and weather

		(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable →		Annual population growth rate, 1970 to 2000					
Independent variables ↓							
Coast/river/topography (7)		No	Yes	Yes	Yes	Yes	Yes
Initial density spline (7)		No	No	Yes	Yes	Yes	Yes
Concentric total pop (7)		No	No	Yes	Yes	Yes	Yes
Ag/mnrl/mnfct (17)		No	No	No	Yes	Yes	Yes
Census divisions (8)		No	No	No	No	Yes	No
Weighted regression		No	No	No	No	No	yes
January daily max temp	Linear	***0.0751***	***0.0663***	***0.0655***	***0.0513***	***0.0497***	***0.0488***
		(0.0098)	***(0.0104)***	***(0.0099)***	***(0.0093)***	***(0.0099)***	***(0.0092)***
	Quadratic	***0.0012***	***0.0012***	***0.0014***	***0.0013***	***0.0016***	***0.0013***
		(0.0004)	***(0.0004)***	***(0.0004)***	***(0.0003)***	***(0.0003)***	***(0.0003)***
July daily heat index	Linear	***−0.0626***	***−0.0508***	***−0.0505***	*−0.0215*	***−0.0242***	−0.0170
		(0.0116)	***(0.0127)***	***(0.0126)***	*(0.0112)*	***(0.0114)***	(0.0114)
	Quadratic	−0.0002	−0.0003	0.0004	−0.0006	−0.0002	−0.0008
		(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
July daily rel humidity	Linear	***−0.0371***	***−0.0410***	***−0.0621***	***−0.0395***	***−0.0549***	***−0.0385***
		(0.0142)	***(0.0147)***	***(0.0162)***	***(0.0129)***	***(0.0132)***	***(0.0123)***
	Quadratic	0.0005	0.0006	−0.0001	−0.0003	***−0.0008***	−0.0003
		(0.0005)	(0.0005)	(0.0004)	(0.0004)	(***0.0003***)	(0.0004)
Annual precipitation	Linear	***0.0216***	***0.0231***	0.0153	−0.0044	−0.0029	−0.0048
		(0.0107)	***(0.0107)***	(0.0100)	(0.0082)	(0.0087)	(0.0080)
	Quadratic	−0.0004	−0.0004	0.0001	0.0002	0.0002	0.0002
		(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Annual precipitation days	Linear	0.0053	0.0041	0.0021	0.0064	0.0061	0.0065
		(0.0060)	(0.0060)	(0.0061)	(0.0048)	(0.0054)	(0.0047)
	Quadratic	***−0.0002***	***−0.0002***	***−0.0003***	***−0.0002***	***−0.0002***	***−0.0002***
		(0.0001)	***(0.0001)***	***(0.0001)***	***(0.0001)***	***(0.0000)***	***(0.0001)***
Observations		3067	3067	3067	3067	3067	3067
# of indep. variables		10	17	31	48	56	48
R-squared		0.272	0.282	0.382	0.503	0.517	0.497
Control variables *R*-squared			0.094	0.226	0.433	0.471	0.423
Marginal *R*-squared			0.188	0.156	0.070	0.046	0.074

Table shows results from regressing ([log(2000 Pop Density) − log(1970 Pop Density)] × 100/30) on the enumerated weather variables, control variables, and a constant. Quadratic weather variables have had their respective sample mean subtracted. Standard errors in parentheses are robust to a spatial correlation using the procedure discussed in the main text. Bold type signifies coefficients that statistically differ from zero at the 0.05 level. Italic type signifies coefficients that statistically different from zero at the 0.10 level. The Column 6 regression weights observations according to 1/(1 + 3000/population).

larger as temperatures increase. This increasing quadratic relationship is intuitive if the advantage of warmer winter weather is the chance to participate in many outdoor recreational activities. Both July heat index and July relative humidity have negative, statistically significant coefficients on their linear terms. Expected population growth falls as summer heat index and relative humidity increase from their respective sample means of 98 °F and 66%.[8]

[8] The negative partial correlation of population growth with summer temperature holds only when controlling for winter temperature. Otherwise, a *positive* coefficient on July heat index statistically differs from zero at the 0.05 level. This statistically-significant positive partial correlation of growth with summer temperature in the absence of a control for winter temperature is quite robust across alternative specifications.

The signs and statistical significance of the coefficients on the winter and summer variables are extremely robust to introducing additional controls. Column 2 includes the coastal proximity and topography variables. Coastal proximity, which helps moderate extreme temperatures, is strongly positively correlated with population growth (Rappaport and Sachs, 2003). But controlling for coastal proximity only slightly tempers the coefficient values on the weather variables.

Column 3 additionally controls for initial density and for total surrounding population in order to capture any growth arising from increasing returns to scale. On the one hand, increasing returns may cause the partial correlation between population and a weather attribute to remain constant even after that attribute's contribution to quality of life and productivity has changed. On the other hand, changes in the degree of increasing returns may induce positive or negative partial correlations between population growth and an attribute notwithstanding unchanged contributions to quality of life and productivity. Directly controlling for the variables from which increasing returns derive should mitigate these problems. Doing so hardly affects the winter and summer coefficients.

One of the explanations for the movement to nice weather is the historical decline of relative U.S. employment in the agricultural, mineral extraction, and heavy manufacturing industries. Column 4 reports results from a regression that includes controls for the initial concentration in each of these. The magnitude of the linear winter coefficient decreases moderately. More substantial decreases in the magnitude of the linear summer coefficients are partly offset by increases in the magnitude of the corresponding quadratic coefficients. The only change in the pattern of statistical significance is that the coefficient on linear heat index differs from zero at the 0.10 rather than the 0.05 level.[9]

As discussed in the Econometric framework section, the controls for initial industrial concentration probably absorb a larger share of variance than is actually *caused* by them. Remaining coefficients would then understate weather's contribution to growth. Many industries require substantial investments that become immobile once they are made. Among industries for which this is so, expanding ones can more easily respond to individuals' shifting locational preferences since they will be investing at a higher rate. As shown below, the move to nice weather began during the 1920s. Thereafter, contracting industries – such as agriculture, mineral extraction, and manufacturing – would become disproportionately concentrated in bad-weather places. Similarly, many of the service-sector industries that came to supplant these contracting industries required less immobile capital. Hence they too would be better able to respond to shifting locational preferences.

Overall, some portion of the move to nice weather probably does reflect the changing industrial composition of employment. But most of it probably does not.

Notwithstanding all the controls in the Column 4 regression, omitted variables might still account for a portion of the partial correlations with winter and summer weather. In case such variables are correlated across regions, Column 5 includes one–zero dummies for 8 of the 9 Census geographic divisions. Coefficient values remain close to those in Column 4. Nor do values change much when dummies for each U.S. state are included (not shown).

Results are similarly robust to a weighted regression that considerably discounts low-population counties. The variance of the error term, $\varepsilon_{i,t}$, in Eq. (3) may be larger for counties with low population. If so, ordinary-least-squares point estimates of parameters will not be efficient.

[9] Results are extremely similar if an expected wage growth variable is additionally included as a control. It is constructed as initial aggregate county labor income in each of 46 industry categories multiplied by the 1970-to-2000 national wage growth of each category (using the BEA REIS dataset). An analogous measure of expected employment growth is not constructed since only 11 industry categories are available to do so.

The optimal GLS estimator involves pre-multiplying the left- and right-hand sides of Eq. (5) by the reciprocal square root of an observation's variance (Davidson and MacKinnon, 1993). Column 6 reports results from such a weighted regression using the same set of controls as in Column 4. The weighting premises that $\varepsilon_{i,t}$ is the sum of two stochastic components: one whose variance is independent of population and one whose variance is proportional to the reciprocal of population. It is assumed that for a county with population of one thousand, 75% of $\varepsilon_{i,t}$ derives from the population-dependent component.[10] This implies that for a county with one hundred residents, 97% of total variance derives from the population-dependent component. For a county with ten thousand residents, 23% of variance derives from it.

The coefficients in the weighted regression (Column 6) are similar to those in the unweighted one (Column 4). The main difference is that the negative coefficient on linear summer heat index no longer statistically differs from zero. But the negative coefficient on July humidity continues to do so at the 0.05 level.[11]

The Column 4, unweighted regression will henceforth serve as a baseline. The magnitude implied by the winter weather coefficients is quite large. An increase in January temperature from one standard deviation below its sample mean to one standard deviation above its sample mean (from 29 °F to 54 °F) is associated with faster growth of 1.3% per year. Miami's Dade County, with an average January daily maximum temperature of 76 °F, has faster expected annual growth of 3.4% per year compared to a county with winter temperature equal to the sample mean. For comparison, the mean population growth rate for U.S. counties from 1970 to 2000 was 0.9% per year with a standard deviation of 1.3%.

The implied magnitudes of the negative coefficients on summer weather in the baseline regression are also large, though slightly less so. An increase in July heat index from one standard deviation below its sample mean to one standard deviation above its sample mean (from 87 °F to 109 °F) is associated with slower growth of 0.5% per year. An increase in relative humidity from one standard deviation below its sample mean to one standard deviation above its sample mean (from 56% to 75%) is associated with slower growth of 0.9% per year.[12] Dade County's average July daily maximum heat index of 109 °F and average July daily relative humidity of 74% imply annual growth 0.7% slower than that of a county with the sample mean values of summer heat index and humidity. Colorado's Summit County, home to many ski resorts, has a mean July heat index of 76 °F and a mean relative humidity of 55%. As a result, it has faster expected annual growth of 0.6% compared to a county with summer weather equal to the sample mean.

In addition to the correlations with summer and winter weather, the regressions in Table 3 show robust evidence of a negative quadratic partial correlation between population growth and precipitation days. The implied magnitude from the Column 4 baseline regression becomes large for places with an especially high number of rainy days. Because of the positive linear coefficient, increasing the number of rainy days by one standard deviation (25 days) above the mean (94 days) leaves expected population growth essentially unchanged. Increasing rainy days by a second and then a third standard deviation slows growth by 0.3 percentage points and then an additional 0.6 percentage points. For Seattle's King County, with an average 182 rainy days per year, annual

[10] Specifically, $E\left(\varepsilon_i^2\right) = \sigma^2\left(1 + \frac{3000}{\text{population}_i}\right)$.

[11] Results are more fragile to a weighting that assumes $\varepsilon_{i,t}$ arises solely from a component with variance proportional to the reciprocal of population.

[12] The reported slower growth from a higher heat index should be interpreted as arising solely from the temperature component. The reported slower growth from an increase in relative humidity includes the effect of the associated rise in heat index. In other words, it holds constant (at its sample mean) only the temperature component of heat index.

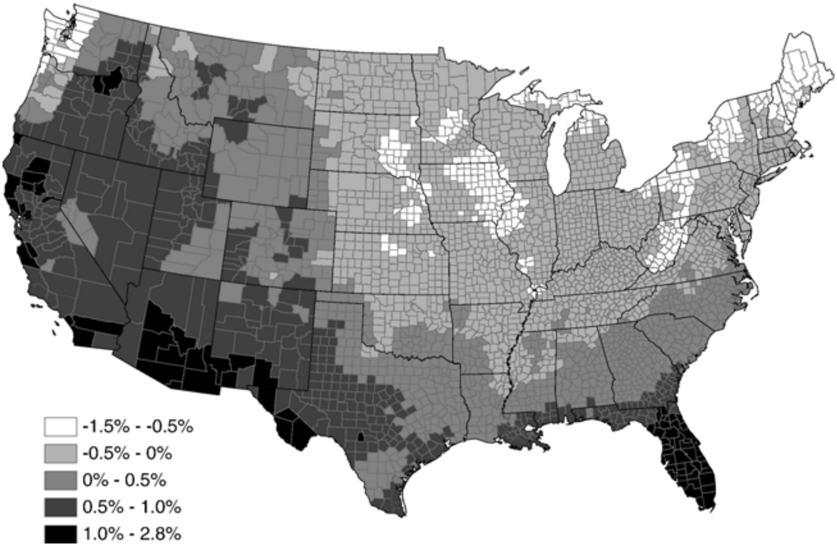

Fig. 2. Expected population growth from weather (1970 to 2000). Figure shows the fitted annual population growth rate controlling for coast, topography, initial density, concentric population, and industrial composition.

expected population growth is 1.3 percentage points slower than that of a county with mean annual precipitation days.

The five weather variables in Table 3 – each entered linearly and quadratically – account for a very large share of the variance of population growth. The *R*-squared value from the regression with no controls is 0.272. For comparison, a regression of population growth on forty-eight continental U.S. state dummies yields an *R*-squared value of 0.322. Additionally including the weather variables with the controls in the baseline Column 4 regression increases *R*-squared by 7.0 percentage points. Again for comparison, the marginal *R*-squared value from adding state dummies to the baseline controls is 12.4 percentage points.

Fig. 2 shows the expected population growth attributable to weather as estimated by the baseline regression.[13] In other words, it shows the vector product of the ten coefficients reported in Column 4 with counties' actual linear and quadratic weather values. The resulting expected growth captures the marginal effect of weather after controlling for coastal proximity, topography, initial density, initial surrounding population, and initial industrial composition. Expected growth from the weather is highest in southern Florida, southwestern Texas, southern New Mexico, southern Arizona, southern California, and coastal northern California. Expected growth from the weather is lowest throughout most of New England and the Midwest as well as in West Virginia and the Pacific Northwest.

The faster expected growth of places with cooler and less humid summers establishes that air conditioning cannot alone account for the move to nice weather. This point is nicely illustrated in Fig. 2 by the high expected growth of California's coastal counties. These counties' mild summer weather implies that their rapid growth cannot be attributed to air conditioning. Conversely, if air conditioning were the main force driving weather-related moves, then expected growth would be much higher throughout the Deep South (one of the hottest and most humid areas of the U.S.).

[13] A color version is included in the paper's supplemental materials.

Nevertheless, the pattern of expected growth suggests that air conditioning was indeed important. The high expected population growth in southern Arizona and southern Florida holds despite average daily summer heat index values exceeding 110 °F. It is hard to imagine this being the case in the absence of air conditioning.

5.2. Alternative growth rates

The movement to nice weather has not been primarily driven by any single demographic group. Rather, growth's positive partial correlation with winter temperature and negative partial correlations with summer heat index and summer humidity hold for a number of sub-populations.

Table 4
Population subgroup growth and weather

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable → Independent variables ↓		Total population growth	Employment growth	Elderly population growth	Working age pop. growth	Immigrant population growth	Native U.S. citizen pop. growth	College graduate pop. growth	College non-graduate pop. growth
Precipitation (4)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coast/river/topography (7)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Initial density spline (7)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Concentric Total Pop (7)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ag/mnrl/mnfct (17)		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
January daily max temp	Linear	***0.0513***	***0.0354***	***0.0607***	***0.0529***	***0.2010***	***0.0439***	***0.0376***	***0.0529***
		(0.0093)	***(0.0098)***	***(0.0084)***	***(0.0097)***	***(0.0234)***	***(0.0093)***	***(0.0121)***	***(0.0090)***
	Quadratic	***0.0013***	***0.0012***	***0.0016***	***0.0012***	−0.0002	***0.0012***	***0.0014***	***0.0013***
		(0.0003)	***(0.0003)***	***(0.0003)***	***(0.0003)***	(0.0009)	***(0.0003)***	***(0.0004)***	***(0.0003)***
July daily heat index	Linear	*−0.0215*	***−0.0248***	***−0.0340***	*−0.0217*	−0.0294	−0.0171	*−0.0253*	−0.0172
		(0.0112)	***(0.0118)***	***(0.0120)***	*(0.0117)*	(0.0299)	(0.0112)	*(0.0152)*	(0.0110)
	Quadratic	−0.0006	−0.0007	***−0.0013***	−0.0002	*−0.0022*	−0.0005	−0.0009	−0.0008
		(0.0005)	(0.0005)	***(0.0005)***	(0.0005)	*(0.0013)*	(0.0005)	(0.0006)	(0.0005)
July daily rel humidity	Linear	***−0.0395***	*−0.0238*	*−0.0224*	***−0.0301***	−0.0499	***−0.0370***	−0.0252	***−0.0378***
		(0.0129)	*(0.0142)*	*(0.0118)*	***(0.0137)***	(0.0383)	***(0.0127)***	(0.0159)	***(0.0128)***
	Quadratic	−0.0003	*−0.0008*	−0.0003	−0.0004	−0.0001	−0.0003	***−0.0013***	0.0000
		(0.0004)	*(0.0004)*	(0.0003)	(0.0004)	(0.0011)	(0.0003)	***(0.0005)***	(0.0003)
Observations		3067	3065	3066	3067	3067	3067	3066	3067
Number of indep. variables		48	48	48	48	48	48	48	48
R-squared		0.503	0.475	0.531	0.507	0.424	0.496	0.327	0.509
Non-weather control variables *R*-squared		0.433	0.445	0.454	0.441	0.284	0.437	0.287	0.419
Marginal *R*-squared		0.070	0.030	0.077	0.066	0.141	0.059	0.040	0.090

Table shows results from regressing annual percentage growth rates for the listed population subgroups on the Column 4 specification in Table 3. Quadratic weather variables have had their respective sample mean subtracted. Standard errors in parentheses are robust to a spatial correlation using the procedure discussed in the main text. Bold type signifies coefficients that statistically differ from zero at the 0.05 level. Italic type signifies coefficients that statistically different from zero at the 0.10 level. For the Column 2 regression, the initial density spline and surrounding total controls are constructed using employment rather than population. For the regressions in Columns 3 and 4, working age is defined to be from 25 to 54; elderly, 65 and above.

Table 4 Column 1 repeats the results from the baseline regression. Column 2 presents results for an analogous regression that uses firms' reported employment growth rather than population growth as its dependent variable. The magnitude of the negative coefficient on linear January temperature is moderately smaller than in the population growth regression, a difference that statistically differs from zero.[14] More important for present purposes is that the negative, statistically significant coefficients on linear and quadratic January temperature imply a quantitatively large, negative partial correlation of employment growth with winter temperature. For summer heat and summer humidity, the implied magnitudes of the negative partial correlations are approximately equal across the two regressions.[15]

The regressions in Columns 3 and 4 partly support the retiree explanation of the move to nice weather. Both the positive correlation with winter temperature and the negative correlation with summer heat index are stronger for elderly population growth (those aged 65 and up) than for working-age population growth (those aged 25 to 54). The differences in the linear coefficients do not statistically differ from zero. But the differences in the quadratic coefficients differ at the 0.10 level for January temperature and at the 0.05 level for July heat index. As above, more important for present purposes is that the coefficients in the working-age regression remain statistically significant and large in magnitude. Elderly migration can account for only part of the overall move to nice weather. Of course, it may be that elderly migration spurred migration of others, for instance by increasing labor demand in destination locales. But the early timing of the move to nice weather, discussed below, casts doubt that this latter possibility was the main driver of the migration.

Another possible explanation for the move to nice weather is that it disproportionately reflects location choices by international immigrants rather than native-born U.S. citizens. The regressions reported in Columns 5 and 6 partly support this. The positive partial correlation with winter temperature is much stronger for immigrants than for natives. Similarly, the implied magnitude of the negative correlation with summer heat index is moderately larger for immigrants. Moreover, the weather variables account for a much higher share of the variation of immigrant growth than of native growth. Even so, growth's positive correlation with winter temperature and negative correlations with heat and humidity remain strong for native-born citizens.

Columns 7 and 8 compare the movement to nice weather between college graduates (i.e., those with at least a bachelor's degree) and non-graduates. The magnitudes of the positive correlation with January temperature and negative correlation with July humidity are moderately larger for the college non-graduates. The magnitude of the negative correlation with July heat index is moderately larger for the college graduates. Once again, both groups are clearly participating in the move to nice weather.

To the extent that the movement to nice weather has been caused by rising incomes, it is intuitive that the same movement would be larger among more highly-educated individuals. After all, such individuals have higher earnings and have seen more rapid increases in earnings. But this intuition misses two important points. First is that the distributional implications of wealth-driven migration are intratemporal rather than intertemporal. An increase in the desirability of nice weather indeed suggests that those at the upper end of the wealth distribution will continue to disproportionately choose to live in nice-weather places. But this static result is perfectly

[14] The statistical significance is based on a regression that has population growth minus employment growth as its dependent variable. This regression and analogous ones for the remaining columns of Table 4 are reported in the paper's supplemental materials.

[15] When the expected wage growth variable discussed in footnote 9 is included in the employment growth regression, negative linear and quadratic coefficients on July heat index both differ from zero at the 0.10 level, but negative linear and quadratic coefficients on July humidity do not.

Table 5
House price growth and weather

		(1)	(2)
Dependent variable → Independent variables ↓		Median house value growth	Median house rent growth
Precipitation (4)		Yes	Yes
Coast/river/topography (7)		Yes	Yes
Initial density spline (7)		Yes	Yes
Concentric total pop (7)		Yes	Yes
Ag/mnrl/mnfct (17)		Yes	Yes
January daily max temp	Linear	***0.0298***	***0.0353***
		(0.0069)	***(0.0038)***
	Quadratic	*−0.0004*	***−0.0003***
		0.0002	***−0.0001***
July daily heat index	Linear	***−0.0353***	***−0.0125***
		(0.0088)	***(0.0043)***
	Quadratic	0.0002	***0.0007***
		(0.0003)	***(0.0002)***
July daily rel humidity	Linear	0.0055	0.0008
		(0.0091)	(0.0056)
	Quadratic	−0.0002	−0.0001
		(0.0003)	(0.0002)
Observations		3063	3049
Number of indep. variables		48	48
R-squared		0.404	0.533
Non-Weather control variables R-squared		0.369	0.362
Marginal R-squared		0.035	0.171

Table shows results from regressing annual percentage growth rates for the listed house price measures on the Column 4 specification in Table 3. Quadratic weather variables have had their respective sample mean subtracted. Standard errors in parentheses are robust to a spatial correlation using the procedure discussed in the main text. Bold type signifies coefficients that statistically differ from zero at the 0.05 level. Italic type signifies coefficients that statistically differ from zero at the 0.10 level.

consistent with a disproportionate increase in the affordability of nice weather by lower-income individuals. Second is that the increase in wages over time can increase the opportunity cost of moving to nice weather. It may be that the associated substitution effect comes closer to offsetting the associated income effect for individuals with more education.

Consistent with compensating differential theory, the migration to nice weather has been paralleled by house price growth. Column 1 of Table 5 reports results from a growth regression of owners' median self-assessed house value on the baseline set of control variables. Column 2 reports results from a similar growth regression of median rental price. In both cases, statistically significant coefficients on linear January temperature and linear July heat index imply that at the respective sample means, house price growth increases as winters become warmer and decreases as summers become hotter. The opposite-signed quadratic coefficients imply that these relationships weaken as temperatures increase. But over observed sample values, the linear coefficients generally dominate the quadratic ones.[16]

[16] Also consistent with compensating differential theory, growth of per-worker labor income has been faster where summer weather is hotter and more humid. It has been approximately orthogonal to winter weather.

Table 6
Population growth and weather by decade

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δlog(Pop Density) → Independent variables ↓		1880–1890	1890–1900	1900–1910	1910–1920	1920–1930	1930–1940	1940–1950	1950–1960	1960–1970	1970–1980	1980–1990	1990–2000
Precipitation		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coast/river/topography		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Initial density spline		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Concentric total pop		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
January daily max temp	Linear	***−0.2061***	***−0.0788***	−0.0322	−0.0056	***0.0698***	***0.0515***	***0.0748***	***0.0682***	***0.0357***	***0.0779***	***0.0579***	***0.0534***
		(0.0490)	***(0.0246)***	(0.0262)	(0.0180)	***(0.0239)***	***(0.0122)***	***(0.0126)***	***(0.0138)***	***(0.0102)***	***(0.0119)***	***(0.0103)***	***(0.0104)***
	Quadratic	−0.0020	***0.0042***	−0.0001	0.0001	***0.0029***	***0.0014***	***0.0017***	***0.0036***	***0.0020***	***0.0015***	***0.0021***	0.0004
		(0.0017)	***(0.0013)***	(0.0011)	(0.0009)	***(0.0010)***	***(0.0005)***	***(0.0004)***	***(0.0006)***	***(0.0005)***	***(0.0005)***	***(0.0004)***	(0.0004)
July daily max heat index	Linear	***0.2756***	***0.0593***	***0.1299***	***0.0431***	−0.0168	***−0.0703***	***−0.0590***	***−0.0531***	***−0.0447***	***−0.0561***	***−0.0520***	***−0.0410***
		(0.0662)	***(0.0276)***	***(0.0320)***	***(0.0219)***	(0.0254)	***(0.0146)***	***(0.0168)***	***(0.0164)***	***(0.0115)***	***(0.0149)***	***(0.0137)***	***(0.0134)***
	Quadratic	***−0.0075***	*−0.0020*	***−0.0034***	***−0.0018***	***−0.0027***	−0.0001	***−0.0026***	***−0.0026***	−0.0003	0.0005	0.0000	*0.0010*
		(0.0022)	*(0.0012)*	***(0.0014)***	***(0.0008)***	***(0.0011)***	(0.0007)	***(0.0007)***	***(0.0007)***	(0.0006)	(0.0007)	(0.0005)	*(0.0005)*
July daily rel humidity	Linear	***0.1301***	*0.0750*	0.0663	*0.0474*	0.0082	0.0158	−0.0313	***−0.0908***	***−0.0833***	***−0.0781***	***−0.0666***	***−0.0391***
		(0.0601)	*(0.0386)*	(0.0422)	*(0.0276)*	(0.0343)	(0.0169)	(0.0193)	***(0.0207)***	***(0.0178)***	***(0.0196)***	***(0.0174)***	***(0.0167)***
	Quadratic	***−0.0071***	0.0007	*−0.0022*	0.0002	***−0.0031***	*0.0009*	−0.0001	***−0.0015***	−0.0009	−0.0004	−0.0003	0.0002
		(0.0015)	(0.0011)	*(0.0012)*	(0.0008)	***(0.0013)***	*(0.0005)*	(0.0006)	***(0.0007)***	(0.0006)	(0.0005)	(0.0005)	(0.0005)
Observations		2395	2604	2696	2844	3014	3060	3062	3064	3063	3067	3067	3069
Independent variables		31	31	31	31	31	31	31	31	31	31	31	31
R-squared		0.735	0.406	0.591	0.122	0.325	0.192	0.368	0.411	0.316	0.311	0.399	0.317
Non-weather control variables *R*-squared		0.664	0.294	0.537	0.104	0.208	0.108	0.283	0.335	0.265	0.183	0.260	0.190
Marginal *R*-squared		0.071	0.112	0.054	0.018	0.118	0.084	0.085	0.076	0.051	0.128	0.138	0.127

Table shows results from regressing annual percentage population growth rates for the listed decade on the Column 3 specification in Table 3. Quadratic weather variables have had their respective sample mean subtracted. Standard errors in parentheses are robust to a spatial correlation using the procedure discussed in the main text. Bold type signifies coefficients that statistically differ from zero at the 0.05 level. Italic type signifies coefficients that statistically differ from zero at the 0.10 level.

5.3. Population growth by decade

The empirical results so far have been for growth rates from 1970 to 2000. But the movement to nice weather actually began in the 1920s. Such a dating reinforces the conclusion that it was largely driven by the broad-based rise in incomes.

Table 6 shows results from decade-by-decade population growth regressions. The specification is the same as in Table 3, Column 3 (the baseline specification excluding the industrial composition controls). Results are qualitatively similar if Census Division dummy variables are additionally included.[17]

From the 1880s through the 1910s, people actually moved to places with bad weather. During the first two decades, population growth was negatively correlated with winter temperature. During all four decades, population growth was positively correlated with summer temperature and humidity. Over the same decades, agriculture's employment share was already declining rapidly. Hence, the shift away from agriculture need not imply movement towards nice weather.

Beginning with the 1920s regression, statistically significant, positive coefficients result on both linear and quadratic January temperature. And statistically significant negative coefficients result on one or more of the summer heat and humidity terms. A first point is that manufacturing's employment share was approximately increasing through 1940 and then approximately constant through 1970. Since the migration towards nice weather persisted through both of these phases, the subsequent decline in manufacturing's employment share seems an unlikely cause. A second point is that the movement towards nice weather far predates the mass adoption of air conditioning, which began in the late 1940s. AC's spread thus seems too late to account for the population shift. Similarly, increases in the life expectancy and financial security of retirees – also largely post World War II phenomena – occurred after the migration towards nice weather was already underway.

Still another explanation for the move to nice weather is that it captures the increasing ease with which people can relocate. Railroads, the automobile, and the airplane have all helped make long-distance moves much less costly. Along with rapidly improving telecommunications technologies, they have also made it much easier to stay in touch with friends and family following long-distance moves. While such falling transportation costs almost certainly helped facilitate the move to nice weather, the timing of the movement suggests that they are unlikely to be the main cause. Railroads were a relatively mature technology many decades before the onset of the move to nice weather. Conversely, it was probably not until the construction of the interstate highway system in the 1950s that the automobile yielded significant time savings on long-distance journeys. Widely affordable air travel arrived even later.

6. Conclusions

Throughout most of the 20th century, U.S. residents migrated to places with nice weather. Possible explanations for this include the introduction of air conditioning, the shift in the industrial composition of U.S. employment, increased elderly migration, and the broad-based rise in incomes. Regressions of 1970-to-2000 population growth on weather suggest that each of these explanations played a role. The positive partial correlations of population growth with

[17] Results with Census Division dummies are reported in the paper's supplemental materials.

summer heat and humidity establish that air conditioning cannot alone account for the migration. But the high expected growth attributable to weather in places with extreme summer heat and humidity almost certainly would not have occurred in the absence of air conditioning. An extensive set of controls for initial industrial structure accounts for a portion of the weather-related moves. But the draw of nice weather nevertheless remains strong. And the move to nice weather has indeed been larger for the elderly, but only moderately so. That the movement to nice weather began in the 1920s reaffirms that it cannot be explained by a combination of air conditioning, shifting industrial structure, and increased elderly migration.

The shortfall of the alternative explanations is consistent with an increasing valuation of weather's contribution to quality of life being an important impetus for the migration. Such an increased valuation is documented in the compensating differential literature and is theoretically predicted from the sharp rise in incomes over the course of the twentieth century. With sufficient complementarity among non-housing goods, housing, and nice weather, the increased valuation of the latter requires migration toward it.

The present set of results suggest several lines of future research. One is the extent to which migration to nice weather has occurred or can be expected to occur in nations and regions other than the United States. Large increases in income have occurred throughout the industrialized world. In many developing countries, income is now rising rapidly or is hoped to do so in the near future. Of course, the degree of labor mobility within many countries is well below that of the U.S. Some initial research along this line finds migration among Japanese prefectures over the period 1955 to 1990 to be negatively correlated with a measure of extreme temperature (Barro and Sala-i-Martin, 1995). Similarly, population growth from 1980 to 2000 *within* EU countries is indeed higher where weather is nicer; but *across* EU nations, population growth is uncorrelated with weather (Cheshire and Magrini, 2004). If European integration continues over time, the lure of better weather may eventually pull people across national borders as well.

A second line of research is the extent to which the U.S. migration to nice weather can be expected to continue. The introduction of air conditioning represents a discrete shock. The transition to the implied new steady-state population distribution might take several decades. But it would eventually end. In contrast, the U.S. population is rapidly aging. And broad-based technological progress and the associated increase in per-capita income are likely to endure. Hence the movement to nice weather may endure as well.

A third line of research is the extent to which migration to places with high levels of other consumption amenities has occurred. Coasts, mountains, lakes, and national parks offer numerous recreational opportunities, as do urban amenities such as restaurants, museums, live performance venues, and professional sports. Other likely draws include low pollution; low traffic; and high-quality schools, universities, and hospitals. It may be that such quality-of-life amenities rather than targeted tax breaks and low wages are the key to future local economic development.

Acknowledgments

Thank you to Jonathan Willis, Steven Durlauf, and several anonymous referees for advice and suggestions. Thank you to Michael Haines for sharing historical census data. Thank you to Taisuke Nakata and Aarti Singh for excellent research assistance. A number of other individuals also contributed to the construction of the underlying data including Nathaniel Baum Snow, Scott Benolkin, Anne Berry, Krista Jacobs, Jason Martinek, Peter Northup, Chris Yenkey, and Andrea Zanter.

References

Ades, Alberto F., Glaeser, Edward L., 1999. Evidence on growth, increasing returns, and the extent of the market. Quarterly Journal of Economics 114 (3), 1025–1045 (Aug).

Barro, Robert J., Sala-i-Martin, Xavier, 1995. Economic Growth. McGraw Hill, New York.

Black, Duncan, Henderson, Vernon J. (2002). "Urban Evolution in the USA." Mimeo, Brown University.

Blomquist, Glenn C., Berger, Mark C., Hoehn, John P., 1988. New estimates of quality of life in urban areas. American Economic Review 78 (1), 89–107 (March).

Cheshire, Paul, Magrini, Stephano, 2004. Population Growth in European Cities: Weather Matters — but only Nationally. Working Paper. London School of Economics/University of Venice.

Ciccone, Antonio, Hall, Robert E., 1996. Productivity and the density of economic activity. American Economic Review 86 (1), 54–70 (March).

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, 2004. Spatial Wage Disparities: Sorting Matters! CEPR Discussion Paper, vol. 4240. February.

Compton, Janice, Pollak, Rober A., 2004. Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas? NBER Working Paper, vol. 10918. November.

Conley, Timothy G., 1999. GMM estimation with cross sectional dependence. Journal of Econometrics 92 (1), 1–45 (Sept.).

Costa, Dora L., Kahn, Matthew E., 2003. The rising price of nonmarket goods. American Economic Review 93 (2), 227–232 (May).

Cragg, Michael I., Kahn, Matthew E., 1999. Climate consumption and climate pricing from 1940 to 1990. Regional Science and Urban Economics 29 (4), 519–539 (July).

Davidson, Russell, MacKinnon, James G., 1993. Estimation and Inference in Econometrics. Oxford University Press, New York.

Fujita, Masahisa, Mori, Tomoya, 1996. The role of ports in the making of major cities: self-agglomeration and hub-effect. Journal of Development Economics 49, 93–120.

Fujita, Masahisa, Thisse, Jacques-Francois, 2002. Economics of Agglomeration — Cities, Industrial Location, and Regional Growth. Cambridge University Press, Cambridge.

Fujita, Masahisa, Krugman, Paul, Venables, Anthony J., 1999. The Spatial Economy. The MIT Press, Cambridge, MA.

Glaeser, Edward L., Kolko, Jed, Saiz, Albert, 2001. Consumer city. Journal of Economic Geography 1 (1), 27–50.

Greenwood, Michael J., Hunt, Gary L., Rickman, Dan S., Treyz, George I., 1991. Migration, regional equilibrium, and the estimation of compensating differentials. American Economic Review 81 (5), 1382–1390 (Dec.).

Gyourko, Joseph, Tracy, Joseph, 1991. The structure of local public finance and the quality of life. Journal of Political Economy 99 (4), 774–806 (Aug.).

Gyourko, Joseph, Tracy, Joseph Kahn, Matthew E. (1999). Handbook of Regional and Urban economics, Volume 3. Cheshire, Paul; Mills, Edwin S., eds. Amsterdam; New York and Oxford:Elsevier Science, North-Holland.

Hanson, Gordon H. (2001). "Market Potential, Increasing Returns, and Geographic Concentration." Mimeo, University of California San Diego (December).

Haurin, Donald R., 1980. The regional distribution of population, migration, and climate. Quarterly Journal of Economics 95 (2), 293–308 (Sept.).

Henderson, J. Vernon, 2003. Marshall's scale economies. Journal of Urban Economics 53 (1), 1–28 (Jan).

Henderson, J. Vernon, Kuncoro, Ari, Turner, Matt, 1995. Industrial development in cities. Journal of Political Economy 103 (5), 1067–1085 (Oct.).

Horan, Patrick M., Hargis, Peggy G. (1995). "County Longitudinal Template, 1840–1990." ICPSR Study No. 6576 [computer file]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. Corrected and amended by Patricia E. Beeson and David N. DeJong, Department of Economics, University of Pittsburgh, 2001. Corrected and amended by Jordan Rappaport, Federal Reserve Bank of Kansas City, 2005.

Krugman, Paul, 1991. Increasing returns and economic geography. Journal of Political Economy 99, 483–499.

Lee, Sanghoon, 2005. Ability Sorting and "Consumer City" Geography. Working Paper. University of British Columbia. Oct.

Maddison, Angus, 1995. Monitoring the World Economy, 1820–1992. Organization for Economic Cooperation and Development, Paris.

Maddison, David, Bigano, Andrea, 2003. The amenity value of the Italian climate. Journal of Environmental Economics and Management 45 (2), 319–332 (March).

Oi, Walter Y., 1997. The Welfare Implications of Invention. In: Bresnahan, Timothy F., Gordon, Robert J. (Eds.), The Economics of New Goods. NBER Studies in Income and Wealth. University of Chicago Press, Chicago.

Rappaport, Jordan, 2004a. Why are population flows so persistent? Journal of Urban Economics 56 (3), 554–580 (November).
Rappaport, Jordan, 2004b. A Simple Model of City Crowdedness. Research Working Paper RWP 04-12. Federal Reserve Bank of Kansas City. December.
Rappaport, Jordan, Sachs, Jeffrey, 2003. The United States as a coastal nation. Journal of Economic Growth 8 (1), 5–46 (March).
Rosen, Sherwin, 1979. Wage-Based Indexes of Urban Quality of Life. In: Miezkowski, Straszheim (Eds.), Current Issues in Urban Economics. Johns Hopkins University Press, Baltimore.
Roback, Jennifer, 1982. Wages, rents, and the quality of life. Journal of Political Economy 90 (6), 1257–1278 (Dec.).
Stover, Mark Erward, Leven, Charles L., 1992. Methodological issues in the determination of the quality of life in urban areas. Urban Studies 29 (5), 737–754.
Stull, Roland B., 2000. Meteorology for Scientists and Engineers, Second Edition. Brooks/Cole, Pacific Grove, CA.
U.S. Bureau for Economic Analysis. "CA05: Personal Income by Major Source and Earnings by Industry" [computer file]. REIS Regional Economic Information System, 1969–2000 CD-Rom (RCN-0295).
U.S. Bureau for Economic Analysis. "CA25: Full-time and Part-time Employment by Industry." [computer file]. REIS Regional Economic Information System, 1969–2000 CD-Rom (RCN-0295).
U.S. Census Bureau. 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. Corrected and amended by Michael R. Haines, Department of Economics, Colgate University, 2001.
U.S. Census Bureau. County and City Data Book, Consolidated File: County Data, 1947–1977, ICPSR Study No. 7736 [computer file]. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 1978. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1981.
U.S. Census Bureau. 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. Census Bureau. 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. Census Bureau. Census of Population and Housing, 2000: Summary File 1 [national computer file]. Internet Release date: November 16, 2001. http://www2.census.gov/census_2000/datasets/%20Summary_File_1/Advance_National.
U.S. Census Bureau. Census of Population and Housing, 2000: Summary File 3 [national computer file]. Internet Release date: November 23, 2002. http://www2.census.gov/census_2000/datasets/%20Summary_File_3/0_National.