CHAPTER

11

EMPIRICAL ANALYSIS OF REGIONAL DATA SETS

A key property of the neoclassical growth model is its prediction of conditional convergence. An economy that starts out proportionately further below its own steady-state position tends to grow faster. We found in Chapters 1 and 2 that economies with similar tastes and technologies converge to the same steady state. In this case, absolute convergence applies; that is, poor economies tend to grow faster than rich ones.

In this chapter, we test the convergence predictions of the neoclassical growth model by looking at the behavior of regions within countries. Although differences in technology, preferences, and institutions do exist across regions, these differences are likely to be smaller than those across countries. Firms and households of different regions within a single country tend to have access to similar technologies and have roughly similar tastes and cultures. Furthermore, the regions share a common central government and therefore have similar institutional setups and legal systems. This relative homogeneity means that absolute convergence is more likely to apply across regions within countries than across countries.

Another consideration in the study of regions is that inputs tend to be more mobile across regions than across countries. Legal, cultural, linguistic, and institutional

barriers to factor movements tend to be smaller across regions within a country than across countries. Hence, the assumption of a closed economy—a standard condition of the neoclassical growth model—is likely to be violated for regional data sets.

We found in Chapter 3 that the dynamic properties of economies that are open to capital movements can be similar to those of closed economies. The key element is that a fraction of the capital stock—which includes human capital—is not mobile, or cannot be used as collateral in interregional or international credit transactions. The speed of convergence is increased by the existence of capital mobility, but remains within a fairly narrow range for reasonable values of the fraction of capital that is mobile. Another result is that a technology without diminishing returns to capital—that is, some version of the AK technology—implies a zero convergence speed whether the economy is open or closed.

We also found in Chapter 9 that the allowance for migration in neoclassical growth models tends to accelerate the process of convergence. The change is, again, a quantitative modification to the speed of convergence. The main point, therefore, is that although regions within a country are relatively open to flows of capital and persons, the neoclassical growth model still provides a useful framework for the empirical analysis.

11.1   TWO CONCEPTS OF CONVERGENCE

We mentioned in Chapter 1 that two concepts of convergence appear in discussions of economic growth across countries or regions. In one view (Barro [1984, Ch. 12], Baumol [1986], DeLong [1988], Barro [1991a], Barro and Sala-i-Martin [1991, 1992a, 1992b]), convergence applies if a poor economy tends to grow faster than a rich one, so that the poor country tends to catch up with the rich one in terms of the level of per capita income or product. This property corresponds to our concept of β convergence.[1] The second concept (Easterlin [1960a], Borts and Stein [1964, Ch.2], Streissler [1979], Barro [1984, Ch. 12], Baumol [1986], Dowrick and Nguyen [1989], Barro and Sala-i-Martin [1991, 1992a, 1992b]) concerns cross-sectional dispersion. In this context, convergence occurs if the dispersion—measured, for example, by the standard deviation of the logarithm of per capita income or product across a group of countries or regions—declines over time. We call this process σ convergence. Convergence of the first kind (poor countries tending to grow faster than rich ones) tends to generate convergence of the second kind (reduced dispersion of per capita income or product), but this process is offset by new disturbances that tend to increase dispersion.

In order to make the relation between the two concepts more precise, we consider a version of the growth equation predicted by the neoclassical growth model

---

[1] This phenomenon is sometimes described as "regression toward the mean."

of Chapter 2. Equation (2.35) relates the growth rate of income per capita between two points in time to the initial level of income. We apply Eq. (2.35) here to discrete periods of unit length (say years) and we also augment it to include a random disturbance:

$$\log(y_{it}/y_{i,t-1}) = a - (1 - e^{-\beta}) \cdot \log(y_{i,t-1}) + u_{it}, \qquad (11.1)$$

where the subscript t denotes the year, and the subscript i denotes the country or region. The theory implies that the intercept, a, equals $x + (1 - e^{-\beta}) \cdot [\log(\hat{y}_i^*)] + x \cdot (t-1)]$, where $\hat{y}_i^*$ is the steady-state level of $\hat{y}_i$. We assume that the random variable u_{it} has 0 mean, variance σ_{ut}^2, and is distributed independently of $\log(y_{i,t-1})$, u_{jt} for $j \neq i$, and lagged disturbances.

We can think of the random disturbance as reflecting unexpected changes in production conditions or preferences. We begin for a single cross section by treating the coefficient a as constant. This specification means that the steady-state value, $\hat{y}_i^*$, and the time trend, $x \cdot (t-1)$, are assumed to be the same for all economies. This assumption is more reasonable for regional data sets than for international data sets; it is plausible that different regions within a country are more similar than different countries with respect to technology and preferences.

If the intercept a is the same in all places and $\beta > 0$, then Eq. (11.1) implies that poor economies tend to grow faster than rich ones. The neoclassical growth models of Chapters 1 and 2 made this prediction. The AK model discussed in Chapter 4 predicts, in contrast, a 0 value for β and, consequently, no convergence of this type. The same conclusion holds for various endogenous growth models (Chapters 6 and 7) that incorporate a linearity in the production function.[2]

Since the coefficient on $\log(y_{i,t-1})$ in Eq. (11.1) is less than 1, the convergence is not strong enough to eliminate the serial correlation in $\log(y_{it})$. Put alternatively, in the absence of random shocks, convergence to the steady state is direct and involves no oscillations or overshooting. Therefore, for a pair of economies, the one that starts out behind is predicted to remain behind at any future date. The models of leapfrogging discussed in Chapter 8 differ in this respect.

Let σ_t^2 be the cross-economy variance of $\log(y_{it})$ at time t. Equation (11.1) and the assumed properties of u_{it} imply that σ_t^2 evolves over time in accordance with the first-order difference equation,[3]

$$\sigma_t^2 = e^{-2\beta} \cdot \sigma_{t-1}^2 + \sigma_{ut}^2, \qquad (11.2)$$

where we have assumed that the cross section is large enough so that the sample variance of $\log(y_{it})$ corresponds to the population variance.

---

[2]We showed, however, in Chapter 4 that β convergence would apply if the technology were asymptotically AK, but featured diminishing returns to capital for finite K.

[3]To derive Eq. (11.2), add $\log(y_{i,t-1})$ to both sides of Eq. (11.1), compute the variance, and use the condition that the covariance between u_{it} and $\log(y_{i,t-1})$ is 0.

If the variance of the disturbance, σ_{ut}^2, is constant over time ($\sigma_{ut}^2 = \sigma_u^2$ for all t), then the solution of the first-order difference Eq. (11.2) is

$$\sigma_t^2 = \frac{\sigma_u^2}{1 - e^{-2\beta}} + \left(\sigma_0^2 - \frac{\sigma_u^2}{1 - e^{-2\beta}}\right) \cdot e^{-2\beta t}, \qquad (11.3)$$

where σ_0^2 is the variance of $\log(y_{i0})$. (It can be readily verified that the solution in Eq. [11.3] satisfies Eq. [11.2].) Equation (11.3) implies that σ_t^2 monotonically approaches its steady-state value, $\sigma^2 = \sigma_u^2/(1 - e^{-2\beta})$, which rises with σ_u^2 but declines with the convergence coefficient, β. Over time, σ_t^2 falls (or rises) if the initial value σ_0^2 is greater than (or less than) the steady-state value, σ^2. Thus, a positive coefficient β (β convergence) does not imply a falling σ_t^2 (σ convergence). To put it another way, β convergence is a necessary but not a sufficient condition for σ convergence.

Figure 11.1 shows the time pattern of σ_t^2 with σ_0^2 above or below σ^2. The convergence coefficient used, $\beta = 0.02$ per year, corresponds to the estimates that we report in a later section. With this value of β, the cross-sectional variance is predicted to fall or rise over time at a slow rate. In particular, if σ_0^2 departs substantially from the steady-state value, σ^2, then it takes about 100 years for σ_t^2 to get close to σ^2.

The cross-sectional dispersion of $\log(y_{it})$ is sensitive to shocks that have a common influence on subgroups of countries or regions. These kinds of disturbances violate the condition that u_{it} in Eq. (11.1) is independent of u_{jt} for $i \neq j$. To the extent that these shocks tend to benefit or hurt regions with high or low income (that is, to the extent that the shocks are correlated with the explanatory variable), the omission of such shocks from the regressions will tend to bias the estimates of β.
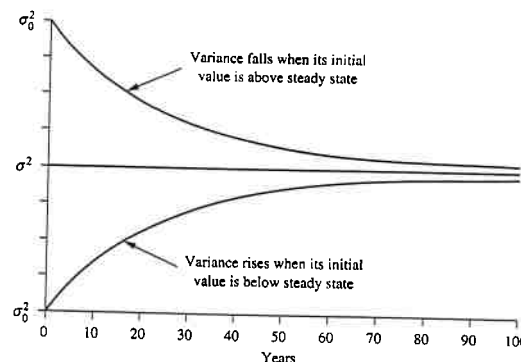

FIGURE 11.1

Theoretical behavior of dispersion. The figure shows the dispersion of per capita product, measured as the variance of the log of per capita product across economies. Although β convergence is assumed to apply, the dispersion may fall, rise, or remain constant, depending on whether it starts above, below, or at its steady-state value, σ^2. The figure assumes $\beta = 0.02$.

Examples are shocks that generate changes in the terms of trade for commodities. For the United States, an example is the sharp drop in the relative prices of agricultural goods during the 1920s. This disturbance had an adverse effect on the incomes of agricultural regions relative to the incomes of industrial regions. We can think also of the two oil price increases of the 1970s and the price decline of the 1980s. These shocks had effects in the same direction on the incomes of regions that produce oil relative to the incomes of other regions. Another example for the United States is the Civil War. This shock had a strong adverse impact on the incomes of southern states relative to the incomes of northern states.

Formally, let S_t be a random variable that represents an economy-wide disturbance for period t. For example, S_t could reflect the relative price of oil as determined on world markets. Then Eq. (11.1) can be modified to

$$\log(y_{it}/y_{i,t-1}) = a - (1 - e^{-\beta}) \cdot \log(y_{i,t-1}) + \varphi_i S_t + u_{it}, \qquad (11.4)$$

where φ_i measures the effect of the aggregate disturbance on the growth rate in region i. If a positive value of S_t signifies an increase in the relative price of oil, then φ_i would be positive for countries or regions that produce a lot of oil.[4] The coefficient φ_i would tend to be negative for economies that produce goods, such as automobiles, that use oil as an input. We think of the coefficient φ_i as distributed cross-sectionally with mean $\bar{\varphi}$ and variance σ_φ^2.

If $\log(y_{i,t-1})$ and φ_i are uncorrelated, then estimates of β in Eq. (11.4) would be unbiased when the shock is omitted from the regression. If $\log(y_{i,t-1})$ and φ_i are positively correlated, then the coefficient estimated by OLS on $\log(y_{i,t-1})$ in Eq. (11.1) would be positively or negatively biased as S_t is positive or negative. As an example, if oil producers have relatively high per capita income, an increase in oil prices will benefit the relatively rich states. Consequently, an OLS regression of growth on initial income will underestimate the true convergence coefficient. In the empirical analysis of the next sections, we hold constant proxies for S_t as an attempt to obtain unbiased estimates of the convergence coefficients.

Equation (11.4) implies that the variance of the log of per capita income evolves as

$$\sigma_t^2 = e^{-2\beta} \cdot \sigma_{t-1}^2 + \sigma_{ut}^2 + S_t^2 \cdot \sigma_\varphi^2 + 2S_t \cdot e^{-\beta} \cdot \text{cov}[\log(y_{i,t-1}), \varphi_i], \quad (11.5)$$

where the variances and covariances are conditioned on the current and past realizations of the aggregate shocks, $S_t, S_{t-1}, \ldots$. If $\text{cov}[\log(y_{i,t-1}), \varphi_i]$ equals 0—that is, if the shock is uncorrelated with initial income—then Eq. (11.5) corresponds to Eq. (11.2), except that realizations of S_t effectively move σ_{ut}^2 around over time. A

---

[4]More precisely, this shock would have a positive effect on the real income derived from the countries or regions that produce a lot of oil. This income may be owned by "foreigners" and appear as part of the net factor payments from "abroad," the term that differentiates GNP from GDP. For example, a substantial fraction of the capital inputs of Wyoming is owned by residents of other states. A positive oil shock will increase Wyoming's nominal GDP (and raise the real value of this GDP when deflated by a national price index) but not necessarily raise its GNP or personal income. For the U.S. states, this distinction is important in a few cases, notably for oil producers.

temporarily large value of S_t raises σ_t^2 above the long-run value σ^2 that corresponds to a typical value of S_t. Therefore, in the absence of a new shock, σ_t^2 returns gradually toward σ^2, as shown in Fig. 11.1.

11.2   CONVERGENCE ACROSS THE U.S. STATES

11.2.1 β Convergence

We now use the data on per capita income for the U.S. states to estimate the speed of convergence, β.[5] (The definitions and sources of the data are in Chapter 10.) Suppose, for the moment, that we have observations at only two points in time, 0 and T. Then Eq. (11.1) implies that the average growth rate over the interval from 0 to T is given by

$$(1/T) \cdot \log(y_{iT}/y_{i0}) = a - [(1 - e^{-\beta T})/T] \cdot \log(y_{i0}) + u_{i0,T}, \qquad (11.6)$$

where $u_{i0,T}$ represents the average of the error terms, u_{it}, between dates 0 and T, and the intercept is $a = x + [(1 - e^{-\beta T})/T] \cdot \log(\hat{y}^*)$.

The coefficient on initial income in Eq. (11.6) is $(1 - e^{-\beta T})/T$, an expression that declines with the length of the interval, T, for a given β. That is, if we estimate a linear relation between the growth rate of income and the log of initial income, then the coefficient is predicted to be smaller the longer the time span over which the growth rate is averaged. The reason is that the growth rate declines as income increases. Hence, if we compute the growth rate over a longer time span, then it combines more of the smaller future growth rates with the initially larger growth rates. Hence, as the interval increases, the effect of the initial position on the average growth rate declines. The coefficient $[(1 - e^{\beta T}/T]$ approaches 0 as T approaches infinity, and it tends to β as T approaches 0. We obtain estimates of β from the nonlinear form of Eq. (11.6), taking account of the value of T that applies in each case. This method should generate similar estimates of β regardless of the length of the averaging interval for the data.

Table 11.1 shows nonlinear least-squares estimates in the form of Eq. (11.6) for 47 or 48 U.S. states or territories for various time periods. The rows of Table 11.1 correspond to different time periods. For example, the first row corresponds to the 110-year period between 1880 and 1990. The first column of the table refers to the equation with only one explanatory variable, the logarithm of income per capita at

---

[5]Barro and Sala-i-Martin (1992) also use the data on Gross State Product (GSP) reported by the Bureau of Economic Analysis. GSP is analogous to GDP in that it assigns the product to the state in which it has been produced. In contrast, income (like GNP) assigns the product to the state in which the owners of the inputs reside. This distinction is potentially important if the economies are open and people tend to own capital in other states, or if there is a lot of interstate commuting (people live in one state and work in another). Barro and Sala-i-Martin (1992) show that, in practice, the distinction turns out not to be that important; the estimates of the speed of convergence for GSP and personal income are similar. Since GSP data are available only starting in 1963, we limit attention in this chapter to the results that use the income data.

TABLE 11.1
Regressions for personal income across U.S. states

	(1)		(2)		(3)	
	Basic equation		Equations with regional dummies		Equations with structural variables & regional dummies	
Period	$\hat{\beta}$	$R^2[\hat{\sigma}]$	$\hat{\beta}$	$R^2[\hat{\sigma}]$	$\hat{\beta}$	$R^2[\hat{\sigma}]$
1880–1990	0.0174 (0.0026)	0.89 [0.0015]	0.0177 (0.0042)	0.93 [0.0012]	—	—
1880–1900	0.0101 (0.0022)	0.36 [0.0068]	0.0224 (0.0040)	0.62 [0.0054]	0.0268 (0.0048)	0.65 [0.0053]
1900–1920	0.0218 (0.0032)	0.62 [0.0065]	0.0209 (0.0062)	0.67 [0.0062]	0.0269 (0.0075)	0.71 [0.0060]
1920–1930	−0.0149 (0.0050)	0.14 [0.0132]	−0.0122 (0.0074)	0.43 [0.0111]	0.0218 (0.0112)	0.64 [0.0089]
1930–1940	0.0141 (0.0030)	0.35 [0.0073]	0.0127 (0.0051)	0.36 [0.0075]	0.0119 (0.0072)	0.46 [0.0071]
1940–1950	0.0431 (0.0049)	0.72 [0.0078]	0.0373 (0.0053)	0.86 [0.0057]	0.0236 (0.0060)	0.89 [0.0053]
1950–1960	0.0190 (0.0035)	0.42 [0.0050]	0.0202 (0.0051)	0.49 [0.0048]	0.0305 (0.0054)	0.66 [0.0041]
1960–1970	0.0246 (0.0040)	0.51 [0.0045]	0.0135 (0.0043)	0.68 [0.0037]	0.0173 (0.0053)	0.72 [0.0036]
1970–1980	0.0198 (0.0063)	0.21 [0.0060]	0.0119 (0.0069)	0.36 [0.0056]	0.0042 (0.0070)	0.46 [0.0052]
1980–1990	0.0011 (0.0100)	0.00 [0.0104]	0.0062 (0.0084)	0.56 [0.0071]	0.0133 (0.0075)	0.75 [0.0055]
Joint, nine subperiods	0.0175 (0.0013)	—	0.0189 (0.0019)	—	0.0220 (0.0021)	—
Likelihood-ratio statistic (p-value)	69.4 (0.000)		31.7 (0.000)		12.6 (0.123)	

Note: The regressions use nonlinear least squares to estimate equations of the form,

$$(1/T) \cdot \log(y_{it}/y_{i,t-T}) = a - [\log(y_{i,t-T})] \cdot [(1 - e^{-\beta T})/T] + \text{other variables},$$

where $y_{i,t-T}$ is per capita income in state i at the beginning of the interval divided by the overall CPI. T is the length of the interval, and the other variables consist of regional dummies and structural measures (see the description in the text). See Chapter 10 for a discussion of the data on the U.S. states. The samples that begin in 1880 have 47 observations. The others have 48 observations. Each column contains the estimate of β, the standard error of this estimate (in parentheses), the R^2 of the regression, and the standard error of the equation (in brackets). The estimated coefficients for constants, regional dummies, and structural variables are not reported. The likelihood-ratio statistic refers to a test of the equality of the coefficients of the log of initial income over the nine subperiods. The p-value comes from a χ^2 distribution with eight degrees of freedom.

the beginning of the period. Column two adds four regional dummies, corresponding to the four main census regions: Northeast, South, Midwest, and West (see Table 10.15 for a list of the states contained in each region). Finally, column three includes sectoral variables that are meant to capture the aggregate shocks discussed in the previous section. We already argued that the inclusion of these auxiliary variables would help to obtain accurate estimates of β.

Each cell contains the estimate of β, the standard error of this estimate (in parentheses), the R^2, and the standard error of the regression (in brackets). All equations have been estimated with constant terms, which are not reported in Table 11.1.

The point estimate of β for the long sample, 1880–1990, is 0.0174 (s.e. = 0.0026).[6] The large R^2, 0.89, can be appreciated from Figure 11.2, which provides a scatter plot of the average growth rate of income per capita between 1880 and 1990 against the log of income per capita in 1880.
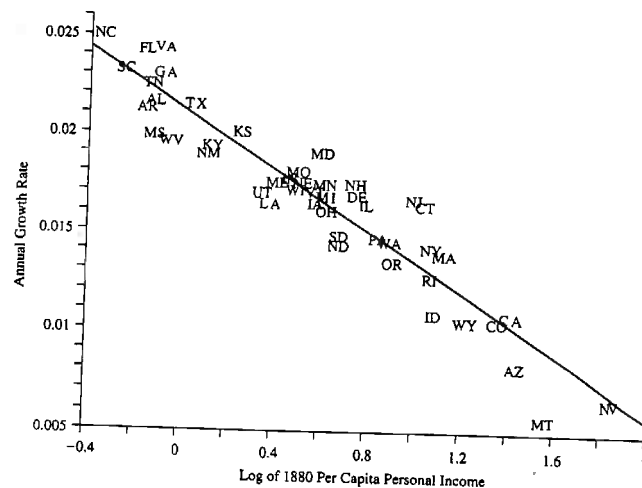

FIGURE 11.2
Convergence of personal income across U.S. states, 1880 personal income and 1880–1990 income growth. The average growth rate of state per capita income for 1880–1990, shown on the vertical axis, is negatively related to the log of per capita income in 1880, shown on the horizontal axis. Thus, absolute β convergence exists for the U.S. states. Each state is represented by its postal code (see Table 10.4).

[6]This regression includes 47 states or territories. Data for the Oklahoma territory are unavailable for 1880.

The second column of the first row presents the estimated speed of convergence when the four regional dummies are incorporated. The estimated β coefficient is 0.0177 (0.0042). The similarity between this estimate and the previous one suggests that the speed at which averages for the four census regions converge is not substantially different from the speed at which averages for the states within each of the regions converge. We can check this result by computing the average income for each of the four regions. The growth rate of a region's average income between 1880 and 1990 is plotted against the log of the region's average income in 1880 in Figure 11.3. The negative relation is clear (the correlation coefficient is -0.97). The estimated speed of convergence implied by this relation is 2.1 percent per year, about the same as the within-region rate shown in column 2.

The next nine rows of Table 11.1 divide the sample into subperiods. The first two are twenty years long (1880 to 1900 and 1920 to 1940), because income data for 1890 and 1910 are unavailable. The remaining seven subperiods are ten years long.

The estimated β coefficient is significantly positive—indicating β convergence—for seven of the nine subperiods. The coefficient has the wrong sign ($\beta < 0$) for only one of the subperiods, 1920–30, a time of large declines in the relative price of agricultural commodities. A likely explanation for this result is that agricultural states tended to be poor states, and the agricultural states suffered the most from
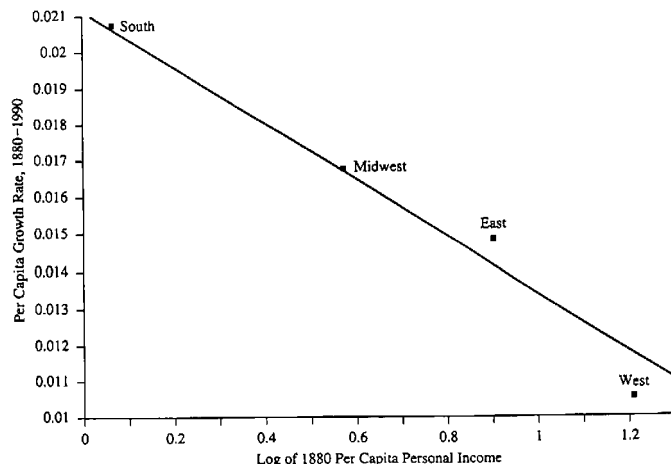

FIGURE 11.3
Convergence of personal income across U.S. regions, 1880 income and 1880–1990 income growth.
The negative relation between income growth and initial income, shown for the U.S. states in Fig. 11.2, applies in Fig. 11.3 to averages over the four main census regions.

the fall in agricultural prices. The estimated coefficient is insignificant for another subperiod, 1980–90, the interval following the oil shocks. (The 1979–81 price increase favored states that were already relatively rich, whereas the 1986 price decline had the opposite effect.)

If we constrain the β coefficients to be the same for all subperiods, then the joint estimate for the basic equation is 0.0175 (0.0013). We reject, however, the hypothesis of stability of the β coefficients over time; the likelihood-ratio statistic is 69.4, with a p-value of 0.000. (Under the null hypothesis, the likelihood-ratio statistic asymptotically follows a χ^2 distribution with eight degrees of freedom.)

Column 2 of Table 11.1 adds regional dummies, where the coefficients of these dummies are allowed to differ for each period. These regional variables capture effects that are common to all states within a region in a given period. The estimated β coefficient for the 1920s still has the wrong sign. Hence, even within regions, poor states tended to grow slower than rich states during the 1920s. The joint estimate for the nine subperiods is now 0.0189 (0.0019), similar to that for the basic regression. We again reject the hypothesis of equality of β coefficients across the subperiods; the likelihood-ratio statistic is now 31.7, with a p-value of 0.000.

Aggregate shocks that affect groups of states differentially, such as shifts in the relative prices of agricultural products or oil, might explain the instability of the estimated coefficients. Following Barro and Sala-i-Martin [1991, 1992a, 1992b], the third column of Table 11.1 adds an additional variable to the regression as an attempt to hold these aggregate shocks constant. The variable, denoted by S_{it} (for structure), is calculated as

$$S_{it} = \sum_{j=1}^{9} \omega_{ij,t-T} \cdot [\log(y_{jt}/y_{j,t-T})/T], \qquad (11.7)$$

where $\omega_{ij,t-T}$ is the weight of sector j in state i's personal income at time $t - T$, and y_{jt} is the national average of personal income per worker in sector j at time t. The nine sectors used are agriculture, mining, construction, manufacturing, trade, finance and real estate, transportation, services, and government. We think of S_{it} as a proxy for the effects reflected in the term $\varphi_i S_t$ in Eq. (11.4).

The structural variable indicates how much a state would grow if each of its sectors grew at the national average rate. For example, suppose that economy i specializes in the production of cars and that the aggregate car sector does not grow over the period between $t - T$ and t. The low value of S_{it} for this region indicates that it should not grow very fast because the car industry has suffered from the shock.

Note from Eq. (11.7) that S_{it} depends on the contemporaneous growth rates of national averages and on lagged values of state i's sectoral shares. For this reason, the variable can be reasonably treated as exogenous to the current growth experience of state i.

Because of lack of data, we can include the structural variable only for the periods after 1929. For the periods before 1929, we obtain a rough measure of S_{it} by using the share of agriculture in the state's total income.

Column three includes structural variables, as well as regional dummies, in the convergence regression. (The coefficients on the regional and structural variables are allowed to differ for each period.) One contrast with the previous results is that the

estimated β coefficient for the 1920s becomes positive. The joint estimate of β for the nine subperiods is 0.022 (0.002). Unlike the previous cases, we now cannot reject at conventional critical levels the hypothesis that the coefficient is the same over the nine subperiods; the likelihood-ratio statistic is 12.6, and the p-value is 0.12.

The main conclusion is that the U.S. states tend to converge at a speed of about two percent per year. Averages for the four census regions converge at a rate that is similar to that for states within regions. If we hold constant measures of structural shocks, then we cannot reject the hypothesis that the speed of convergence is stable over time.

11.2.2 Measurement Error

The existence of temporary measurement error in income tends to introduce an upward bias in the estimate of β; that is, the elimination of measurement error over time can generate the appearance of convergence.[7] One reason for measurement error is that each state's nominal income is deflated by a national price index, because accurate indexes do not exist at the state level.

One approach to handle measurement error is to use earlier lags of the log of income as instruments in the regressions. If measurement error is temporary (and the error term is not serially correlated), then the earlier lags of the log of income would be satisfactory instruments for the log of income at the start of each period. If we reestimate column 1 of Table 11.1 with the previous lag of the log of income used as an instrument, then we get a joint estimate of β of 0.0198 (0.0016). This panel uses eight subperiods starting in 1900 because the observation for 1880–1900 is lost. The OLS estimate of β for the same eight subperiods is 0.0202 (0.0015). Hence, the use of instruments generates a minor decline in the estimate of β.

When we estimate the subperiods separately, we again find only a small difference between the instrumental-variable (IV) and OLS estimates. The largest change applies to the period 1950–60, for which the IV estimate is 0.0156 (0.0036), compared with the OLS value of 0.0190 (0.0036).

The results for columns 2 and 3 of Table 11.1 are similar. For example, for column 3, the joint IV estimate is 0.0197 (0.0026), compared with an OLS estimate of 0.0206 (0.0024). Thus, the main finding is that the instrumental procedures do not alter the basic findings on β convergence. Our conclusion is that measurement error is unlikely to be a key element in the results.

11.2.3 σ Convergence

Figure 11.4 shows the cross-sectional standard deviation for the log of per capita personal income net of transfers for 47 or 48 U.S. states or territories from 1880 to 1992. The dispersion declined from 0.54 in 1880 to 0.33 in 1920, but then rose to

---

[7]The same property holds for short-term business fluctuations. We may want to design a model in which these temporary fluctuations of output are distinguished from the kinds of transitional dynamics that appear in growth models.
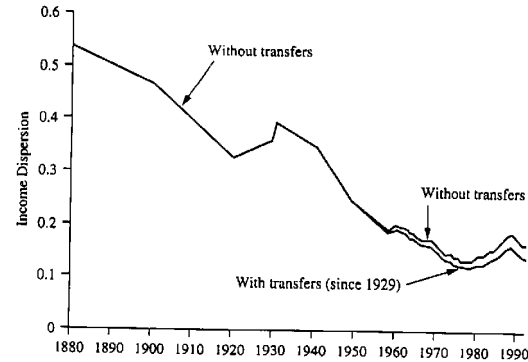
FIGURE 11.4
Dispersion of personal income across U.S. states, 1880–1992. The figure shows the cross-sectional variance of the log of per capita personal income for 47 or 48 U.S. states or territories from 1880 to 1992. This measure of dispersion declined from 1880 to 1920, rose in the 1920s, fell from 1930 to the mid-1970s, rose through 1988, and declined again through 1992.

0.40 in 1930. This rise reflects the adverse shock to agriculture during the 1920s; the agricultural states were relatively poor in 1920 and suffered a further reduction in income with the fall in agricultural prices.

After 1930, the dispersion fell to 0.35 in 1940, 0.24 in 1950, 0.21 in 1960, 0.17 in 1970, and a low point of 0.14 in 1976. The long-run decline stopped in the mid-1970s, after the first oil shock, and σ_t rose to 0.15 in 1980 and 0.19 in 1988. The rise in dispersion was reversed at the end of the 1980s (apparently as soon as Mr. Reagan was no longer President), and dispersion fell through 1992.

Figure 11.4 shows that the behavior of the cross-sectional dispersion of personal income net of transfers is similar to that of gross income of transfers. In particular, dispersion fell for both concepts of income after 1930, rose between 1977 and 1988, and fell between 1988 and 1992. Although the transfers tend to reduce the cross-state dispersion of per capita income, changes in the amount of transfers are not the main source of the long-run decline in income dispersion across the states.

11.3 CONVERGENCE ACROSS JAPANESE PREFECTURES

11.3.1 β Convergence

We now analyze the pattern of β convergence for per capita income across 47 Japanese prefectures, using the data set of Barro and Sala-i-Martin (1992b). (See Chapter 10 for the sources and definitions.) Table 11.2 reports nonlinear estimates of the convergence coefficient, β, for the period 1930–90. The setup of Table 11.2 parallels that of Table 11.1.

TABLE 11.2
Regressions for personal income across Japanese prefectures

Period	(1) Basic equation		(2) Equations with district dummies		(3) Equations with structural variables & district dummies	
	$\hat{\beta}$	$R^2[\hat{\sigma}]$	$\hat{\beta}$	$R^2[\hat{\sigma}]$	$\hat{\beta}$	$R^2[\hat{\sigma}]$
1930–90	0.0279 (0.0033)	0.92 [0.0019]	0.0276 (0.0024)	0.97 [0.0012]	—	—
1930–55	0.0358 (0.0035)	0.86 [0.0045]	0.0380 (0.0037)	0.90 [0.0038]	—	—
1955–90	0.0191 (0.0035)	0.59 [0.0027]	0.0222 (0.0035)	0.81 [0.0020]	—	—
1955–1960	-0.0152 (0.0079)	0.07 [0.0133]	-0.0023 (0.0082)	0.44 [0.0111]	0.0047 (0.0118)	0.46 [0.0112]
1960–1965	0.0296 (0.0072)	0.30 [0.0108]	0.0360 (0.0079)	0.55 [0.0093]	0.0414 (0.0096)	0.56 [0.0093]
1965–1970	-0.0010 (0.0062)	0.00 [0.0097]	0.0127 (0.0067)	0.47 [0.0076]	0.0382 (0.0091)	0.62 [0.0065]
1970–1975	0.0967 (0.0100)	0.78 [0.0095]	0.0625 (0.0092)	0.87 [0.0078]	0.0661 (0.0118)	0.87 [0.0079]
1975–1980	0.0338 (0.0100)	0.23 [0.0087]	0.0455 (0.0119)	0.37 [0.0085]	0.0469 (0.0145)	0.37 [0.0086]
1980–1985	-0.0115 (0.0077)	0.04 [0.0075]	0.0076 (0.0089)	0.37 [0.0066]	0.0102 (0.0094)	0.37 [0.0067]
1985–1990	0.0007 (0.0067)	0.00 [0.0067]	0.0086 (0.0082)	0.28 [0.0061]	0.0085 (0.0085)	0.28 [0.0062]
Joint, seven subperiods	0.0125 (0.0032)	—	0.0232 (0.0034)	—	0.0312 (0.0040)	—
Likelihood-ratio statistic (p-value)	94.6 (0.000)		40.6 (0.000)		26.4 (0.002)	

Note: See Chapter 10 for a discussion of the data on Japanese prefectures, and see the note to Table 11.1 for the form of the regressions. The variable $y_{i,t-T}$ is per capita income in prefecture i at the beginning of the interval divided by the overall CPI. All samples have 47 observations. The likelihood-ratio statistic refers to a test of the equality of the coefficients of the log of initial income over the seven subperiods. The p-value comes from a χ^2 distribution with six degrees of freedom.

The first row of Table 11.2 pertains to regressions for the whole period, 1930–90. The basic equation in column 1 includes only the log of initial income as a regressor. The estimated β coefficient is 0.0279 (0.0033), with an R^2 of 0.92. The good fit can be appreciated in Fig. 11.5. The strong negative correlation between the growth rate from 1930 to 1990 and the log of per capita income in 1930 confirms the existence of β convergence across the Japanese prefectures.

The estimated β coefficient is essentially the same in column 2, which incorporates dummies for the seven Japanese districts as explanatory variables. (See the notes to Table 10.18 for a listing of the prefectures contained in each district.) This
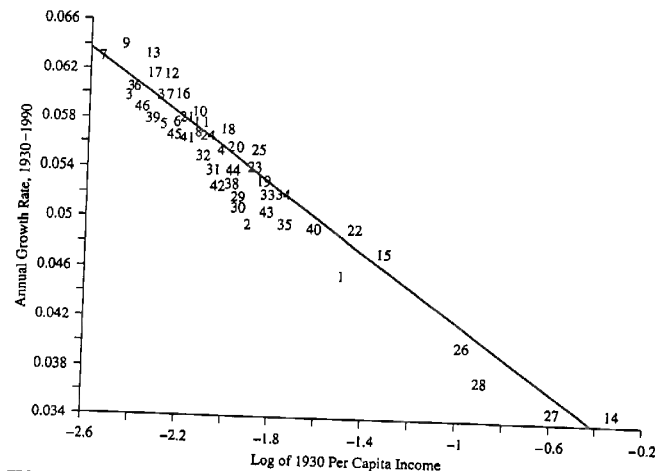
FIGURE 11.5
Convergence of personal income across Japanese prefectures, 1930 income and 1930–1990 income growth. The average growth rate of prefectural per capita income for 1930–90, shown on the vertical axis, is negatively related to the log of per capita income in 1930, shown on the horizontal axis. Thus, absolute β convergence exists for the Japanese prefectures. The numbers shown identify the prefecture (see Table 10.7).

finding suggests that the speed of convergence for prefectures within districts is similar to that across districts. This idea can be checked by running a regression that uses the seven data points for the growth and level of the average per capita income of districts. The negative relation between the growth rate from 1930 to 1990 and the log of per capita income in 1930 is displayed in Fig. 11.6. The β coefficient estimated from these observations is 0.0261 (0.0079). Hence, we confirm that the speed of convergence across districts is about the same as that within districts.

The second and third rows of Table 11.2 break the full sample into two long subperiods, 1930–55 and 1955–90. For the basic equation, the speed of convergence for the first subperiod is larger than that for the second, 0.0358 (0.0035) versus 0.0191 (0.0035). The same relation holds for the second column, which adds the district dummies as explanatory variables. (Different coefficients on the dummies are estimated for the two subperiods.) Hence, we conclude that the speed of convergence after 1955 was substantially slower than that between 1930 and 1955. The lack of sectoral data for the early period does not, however, allow us to investigate the cause of this difference. We therefore restrict the rest of the analysis to the post-1955 period.

The next seven rows of Table 11.2 break the sample into five-year subperiods, starting in 1955. For three of the subperiods, the sign of the estimated β coefficient in the basic equation is opposite to the one expected. The speed of convergence is
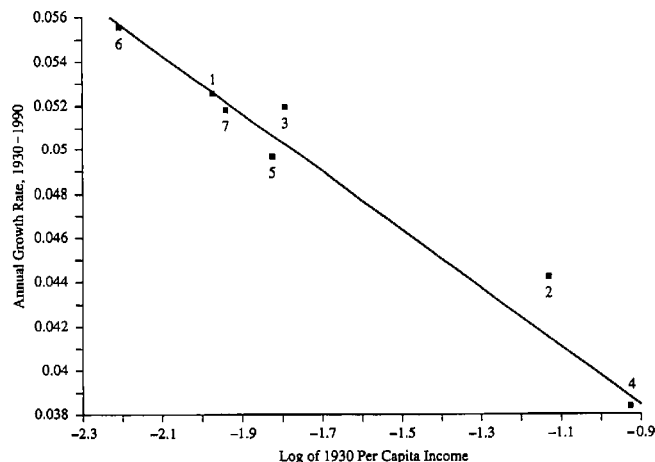
FIGURE 11.6
Convergence of personal income across Japanese districts, 1930 income and 1930-1990 income growth. The negative relation between income growth and initial income, shown for Japanese prefectures in Fig. 11.5, applies also in Fig. 11.6 to averages for the seven major districts.

positive and significant for the periods 1960-65, 1970-75, and 1975-80. The joint estimate for the seven subperiods is 0.0125 (0.0032). A test for the equality of coefficients over time is strongly rejected; the likelihood-ratio statistic is 94.6, with a p-value of 0.000.

The results with district dummies in column 2 allow for different coefficients on the dummies in each subperiod. In this case, only the estimated β coefficient for 1955-60 has the wrong sign, and it is not significant. The joint estimate is 0.0232 (0.0034). However, we still reject the equality of coefficients at the 5 percent level; the likelihood-ratio statistic is 40.6, with a p-value of 0.000.

Column 3 adds a measure of the structural variable, S_{it}, defined in Eq. (11.7). This variable is analogous to the one constructed for the U.S. states. The coefficients on the structural variable are allowed to differ for each subperiod. In contrast with the previous two columns, none of the subperiods has the wrong sign when the sectoral variable is included. The joint estimate for the seven subperiods is 0.0312 (0.0040). The likelihood-ratio statistic for the equality of coefficients over time is 26.4, with a p-value of 0.002. Therefore, although the likelihood-ratio statistic is smaller than before, we still reject the hypothesis of stability over time.

One source of instability in the estimated β coefficients is that Tokyo is an outlier in the 1980s: Tokyo was by far the richest prefecture in its district in 1980 and had the largest growth rate from 1980 to 1990, an outcome not captured by the

structural variable that we have included. If we add a dummy for Tokyo for the 1980s, then we get estimated β coefficients of 0.0218 (0.0112) for 1980-85 and 0.0203 (0.0096) for 1985-90. With this dummy included, the test of equality of coefficients over the seven subperiods yields a likelihood-ratio statistic of 16.7, which implies a p-value of 0.010.

Another source of instability is the period 1970-75, for which the estimated β coefficient of 0.0661 (0.0118) is substantially higher than the others. A likely explanation for this high estimated value of β is that the oil shock of 1973 had an especially adverse impact on the richer industrial areas. The structural variable is supposed to hold constant this type of shock, but the construct that we have been able to measure does not seem to capture this effect.

As with the U.S. states, we reestimated the equations for Japanese prefectures with earlier lags of income used as instruments. The conclusion again is that the estimates are not materially affected. For example, for column 3 of Table 11.2, the joint estimate of β falls from 0.0312 (0.0040) to 0.0282 (0.0042) when the instruments are used.

11.3.2 σ Convergence across Prefectures

We want now to assess the extent to which there has been σ convergence across prefectures in Japan. We calculate the unweighted cross-sectional standard deviation for the log of per capita income, σ_t, for the 47 prefectures from 1930 to 1990. Figure 11.7 shows that the dispersion of personal income increased from 0.47 in 1930 to 0.63 in 1940. One explanation of this phenomenon is the explosion of military spending during the period. The average growth rates for districts 1 (Hokkaido-Tohoku) and
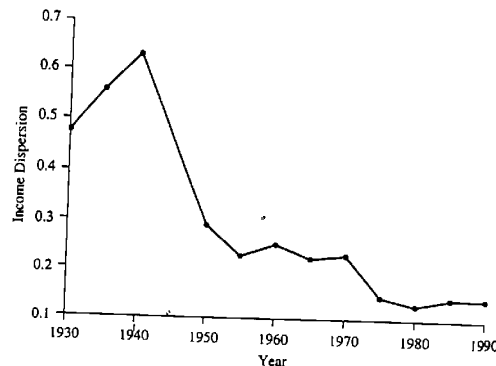

FIGURE 11.7
Dispersion of personal income across Japanese prefectures, 1930-1990. The figure shows the cross-sectional variance of the log of per capita personal income for 47 Japanese prefectures from 1930 to 1990. This measure of dispersion fell from the end of World War II until 1980.

7 (Kyushu), which are mainly agricultural, were −2.4 percent and −1.7 percent per year, respectively. In contrast, the industrial regions of Tokyo, Osaka, and Aichi grew at 3.7, 3.1, and 1.7 percent per year, respectively.

The cross-prefectural dispersion decreased dramatically after World War II: it fell to 0.29 in 1950, 0.25 in 1960, 0.23 in 1970, and hit a minimum of 0.12 in 1978. The dispersion increased slightly then: σ_t rose to 0.13 in 1980, 0.14 in 1985, and 0.15 in 1987, but has been relatively stable since 1987. Thus, the pattern is similar to that for the U.S. states.

11.4 CONVERGENCE ACROSS EUROPEAN REGIONS

11.4.1 β Convergence

We now analyze convergence for 90 regions in eight European countries: 11 in Germany, 11 in the United Kingdom, 20 in Italy, 21 in France, 4 in the Netherlands, 3 in Belgium, 3 in Denmark, and 17 in Spain. The data, described in Chapter 10, correspond to GDP per capita for the first seven countries and to income per capita for Spain.

Table 11.3 shows the estimates of β in the form of Eq. (11.6) for the period 1950–90. The regressions include country dummies for each period to proxy for

TABLE 11.3
Convergence across European regions

Period	(1) Equations with country dummies		(2) Equations with sectoral shares and country dummies	
	$\hat{\beta}$	$R^2[\hat{\sigma}]$	$\hat{\beta}$	$R^2[\hat{\sigma}]$
1950–60	0.018 (0.006)	0.83 [0.0099]	0.034 (0.009)	0.84 [0.0094]
1960–70	0.023 (0.009)	0.97 [0.0065]	0.020 (0.006)	0.97 [0.0064]
1970–80	0.020 (0.009)	0.99 [0.0079]	0.022 (0.007)	0.99 [0.0077]
1980–90	0.010 (0.004)	0.97 [0.0066]	0.007 (0.005)	0.97 [0.0064]
Joint, four subperiods	0.019 (0.002)	— —	0.018 (0.003)	—
Likelihood-ratio statistic (p-value)	4.9 (0.179)		8.6 (0.034)	

Note: See Chapter 10 for a discussion of the data on European regions, and see the note to Table 11.1 for the form of the regressions. The variable $y_{i,t-T}$ is an index of the per capita GDP (income for Spain) in region i at the beginning of the interval. All samples have 90 observations. The likelihood-ratio statistic refers to a test of the equality of the coefficients of the log of initial per capita GDP or income over the four subperiods. The p-value comes from a χ^2 distribution with three degrees of freedom.

differences in the steady-state values of x_i and $\hat{y}_i^*$ in Eq. (11.1) and for countrywide fixed effects in the error terms. The country dummies, which are not reported in Table 11.3, have substantial explanatory power. The first four rows of column 1 show the results for four decades. The estimates of β are reasonably stable over time; they range from 0.010 (0.004) for the 1980s to 0.023 (0.009) for the 1960s. The joint estimate for the four decades is 0.019 (0.002). The hypothesis of constant β over time cannot be rejected at conventional levels of significance; the likelihood ratio statistic is 4.9, with a p-value of 0.18.

Figure 11.8 shows for the 90 regions the relation of the growth rate of per capita GDP (income for Spain) from 1950 to 1990 (1955 to 1987 for Spain) to the log of per capita GDP or income at the start of the period. The variables are measured relative to the means of the respective countries. The figure shows the negative relation that is familiar from the U.S. states and Japanese prefectures. The correlation between the growth rate and the log of initial per capita GDP or income in Fig. 11.8 is −0.72. Since the underlying numbers are expressed relative to own-country means, the relation in Fig. 11.8 pertains to β convergence within countries, rather than between countries. The graph therefore corresponds to the estimates that include country dummies in column 1 of Table 11.3.
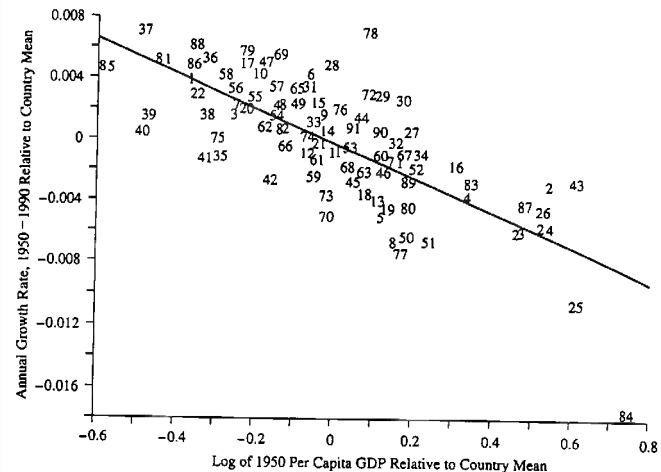

FIGURE 11.8
Growth rate from 1950 to 1990 versus 1950 per capita GDP for 90 regions in Europe. The growth rate of a region's per capita GDP for 1950–90, shown on the vertical axis, is negatively related to the log of per capita GDP in 1950, shown on the horizontal axis. The growth rate and level of per capita GDP are measured relative to the country means. Hence, this figure shows that absolute β convergence exists for the regions within Germany, the United Kingdom, Italy, France, the Netherlands, Belgium, Denmark, and Spain. The numbers shown identify the regions (see Table 10.5).

Column 2 adds the share of agriculture and industry in total employment or GDP at the start of each subperiod.[8] These share variables are as close as we can come with our present data for the European regions to measuring the structural variable, S_{it}, that appears in Eq. (11.7). The results allow for period-specific coefficients for the sectoral shares.

The joint estimate of β for the four subperiods is now 0.018 (0.003). The test of the hypothesis of stability of β across periods yields a likelihood-ratio statistic of 8.6, with a p-value of 0.034. Thus, in contrast to our findings for the United States and Japan, the inclusion of the share variables makes the β coefficients appear less stable over time. Probably, a better measure of structural composition would yield more satisfactory results.

We have also estimated the joint system for Europe with individual β coefficients for the five large countries (Germany, the United Kingdom, Italy, France, and Spain). This system corresponds to the four-period regression shown in column 2 of Table 11.3, except that the coefficient β is allowed to vary over the countries (but not over the subperiods). This system contains country dummies (with different coefficients for each subperiod) and share variables (with coefficients that vary over the subperiods but not across the countries). The resulting estimates of β are Germany (11 regions), 0.0224 (0.0067); United Kingdom (11 regions), 0.0277 (0.0104); Italy (20 regions), 0.0155 (0.0037); France (21 regions), 0.0121 (0.0061); and Spain (17 regions), 0.0182 (0.0048). Note that the individual point estimates are all close to 2 percent per year; they range from 1.2 percent per year for France to 2.8 percent per year for the United Kingdom.

The likelihood-ratio statistic for equality of the β coefficients across the five countries is 3.0, and the corresponding p-value is 0.55. Hence, we cannot reject the hypothesis that the speed of regional convergence within the five European countries is the same.

We also reestimated the European equations with earlier lags of per capita GDP or income used as instruments. This procedure necessitated the elimination of the first subperiod; hence, only the three decades from 1960 to 1990 are now included. The use of instruments had little impact on the results that included only country dummies, corresponding to column 1 of Table 11.3. The joint estimate of β goes from 0.0187 (0.0022) in the OLS case (with only three subperiods included) to 0.0165 (0.0023). If the agricultural and industrial share variables are added, however, the joint estimate of β goes from 0.0153 (0.0034) to 0.0073 (0.0038). We think that the sharp drop in the estimated β coefficient in this case reflects inadequacies in the share variables as measures of structural shifts.

11.4.2 σ Convergence

Figure 11.9 shows the behavior of σ_t for the regions within the five large countries: Germany, the United Kingdom, Italy, France, and Spain. The countries are always
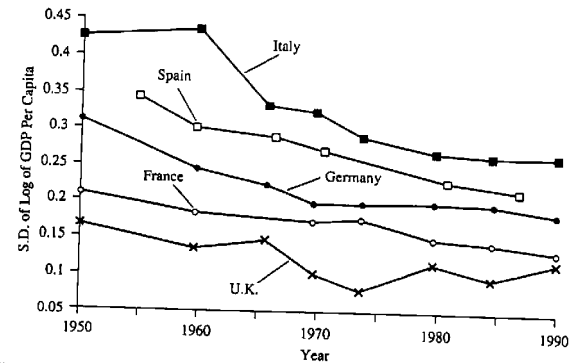
FIGURE 11.9
Dispersion of per capita GDP within five European countries. The figure shows the cross-sectional variance of the log of per capita GDP from 1950 to 1990 for 11 regions in Germany, 11 in the United Kingdom, 20 in Italy, 21 in France, and 17 in Spain. This measure of dispersion fell in most cases since 1950, but has been roughly stable in Germany and the United Kingdom since 1970.

ranked in descending order of dispersion as Italy, Spain, Germany, France, and the United Kingdom. The overall pattern shows declines in σ_t over time for each country, although little net change occurs since 1970 for Germany and the United Kingdom. The rise in σ_t from 1974 to 1980 for the United Kingdom—the only oil producer in the European sample—likely reflects the effect of oil shocks. In 1990, the values of σ_t are 0.27 for Italy, 0.22 for Spain (for 1987), 0.19 for Germany, 0.14 for France, and 0.12 for the United Kingdom.

11.5 MIGRATION ACROSS THE U.S. STATES

This section considers the empirical determinants of net migration among the U.S. states. The analysis in Section 9.1.3 suggests that m_{it}, the annual rate of net migration into region i between years $t - T$ and t, can be described by a function of the form

$$m_{it} = f(y_{i,t-T}, \theta_i, \pi_{i,t-T}; \text{variables that depend on } t \text{ but not } i), \quad (11.8)$$

where $y_{i,t-T}$ is per capita income at the beginning of the period, θ_i is a vector of fixed amenities (such as climate and geography), and $\pi_{i,t-T}$ is the population density in region i at the beginning of the period.[9] The set of variables that depends on t

---

[8]The share figures for the first three subperiods are based on employment. The values for 1980–90 are based on GDP.

[9]Some amenities, such as government policies with respect to tax rates and regulations, would vary over time. We do not deal with these types of variables in the present analysis.

but not on i includes any elements that influence per capita incomes and population densities in other economies. Also included are effects like technological progress in heating and air conditioning, changes that alter people's attitudes about weather and population density.

Per capita income—a proxy for wage rates—would have a positive effect on migration, whereas population density would have a negative effect. The functional form that we implement empirically is

$$m_{it} = a + b \cdot \log(y_{i,t-T}) + c_1\theta_i + c_2\pi_{i,t-T} + c_3 \cdot (\pi_{i,t-T})^2 + v_{it}, \quad (11.9)$$

where v_{it} is an error term, $b > 0$, and the form allows for a quadratic in population density, $\pi_{i,t-T}$. The marginal effect of $\pi_{i,t-T}$ on m_{it} is negative if $c_2 + 2c_3 < 0$.

Although there is an extensive literature about variables to include as amenities, θ_i, the present analysis includes only the log of average heating-degree days, denoted $\log(HEAT_i)$, which is a disamenity so that $c_1 < 0$. The variable $\log(HEAT_i)$ has a good deal of explanatory power for net migration across the U.S. states. We considered alternative measures of the weather, but they did not fit as well. It would be useful to include migration for retirement, a mechanism that likely explains outliers such as Florida. However, these kinds of modifications probably would not change the basic findings that we now present about the relation between net migration and state per capita income.

Our data on net migration for the U.S. states start in 1900 and are available for every census year except 1910 and 1930. Chapter 10 describes the sources and definitions of these data. We calculate the ten-year annual migration rates into a state by dividing the number of net migrants between dates $t - T$ and t by the state's population at date $t - T$.

Figure 11.10 shows the simple long-term relation between the migration rate and the log of initial income per capita.[10] The horizontal axis plots the log of state per capita income in 1900. The positive association is evident (correlation = 0.51). The main outlier is Florida, which has a lower than average initial income per capita and a very high net migration rate of 3 percent per year.

Table 11.4 shows regression results in the form of Eq. (11.9) for net migration into U.S. states. The results reported are for eight subperiods starting with 1900–20. The regressions include period-specific coefficients for $\log(y_{i,t-T})$ and for the log of heating-degree days. (The hypothesis of stability over the subperiods in the coefficients of $\log[HEAT_i]$ is rejected at the 5 percent level, although the estimated coefficients on $\log[y_{i,t-T}]$ change little if only a single coefficient is estimated for the heat variable.) Since the hypothesis that the coefficients for the population-density variables are stable over time is accepted at the 5 percent level, we estimate Eq. (11.9) with one coefficient for the density and one for the square of the density. The regressions also include period-specific coefficients for regional dummies and structural-share variables. (The estimated coefficients for the regional and structural variables are sometimes significant, but play a minor role overall.)

[10]The variable on the vertical axis is the average annual in-migration rate for each state from 1900 to 1987. The variable is the average for each subperiod weighted by the length of the interval.
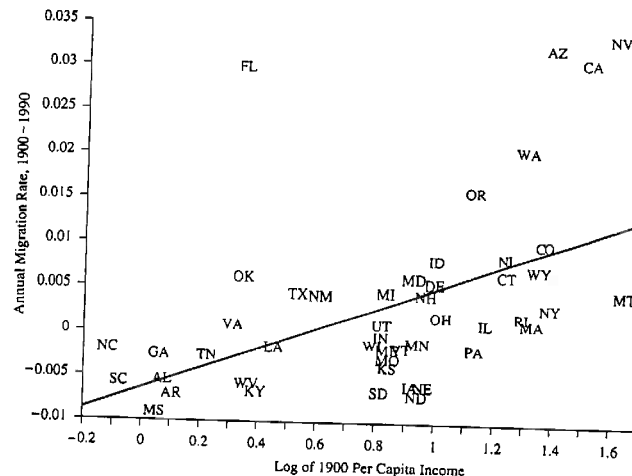
FIGURE 11.10

Migration and initial state income, 1900–1990. The average net migration rate for 48 U.S. states or territories from 1900 to 1990, shown on the vertical axis, is positively related to the log of initial per capita income, shown on the horizontal axis. Florida, Arizona, California, and Nevada have notably higher net migration rates than the values predicted by their initial levels of income.

The estimated coefficients for $\log(HEAT_i)$ in Table 11.4 are all negative and most are significantly different from 0; other things equal, people prefer warmer states. The jointly estimated coefficients for density are -0.043 (0.008) on the linear term and 0.030 (0.010) on the squared term. These point estimates imply that the marginal effect of population density on migration is negative for all states, except for the three with the highest densities: New Jersey, Rhode Island since 1960, and Massachusetts since 1970.

The coefficient on the log of initial per capita income is significantly positive for all subperiods. The joint estimate is 0.0260 (0.0023), which implies a t-value over 11. The estimated response of migration to the log of initial level is, however, not stable over time; the likelihood-ratio statistic is 17.1, with a p-value of 0.017. The main sources of instability are the unusually large coefficients on income in the 1950s and 1960s; the coefficients in these two subperiods are 0.0438 (0.0086) and 0.0435 (0.0083), respectively.

Although highly significant, the coefficient on initial income, 0.026, is small in an economic sense. The coefficient means that, other things equal, a 10 percent differential in income per capita raises net in-migration only by enough to raise the area's rate of population growth by 0.26 percent per year. Our previous results suggest that differences in per capita income tend themselves to vanish at a slow speed,

TABLE 11.4
Regressions for net migration into U.S. states, 1900–89

Period	Log of per capita income	Heating-degree days	Population density	Square of population density	$R^2[\hat{\sigma}]$
1900–20	0.0335 (0.0075)	−0.0066 (0.0037)	−0.0433 (0.0079)	0.0307 (0.0095)	0.70 [0.0111]
1920–30	0.0363 (0.0078)	−0.0124 (0.0027)	−0.0433 (0.0079)	0.0307 (0.0095)	0.61 [0.0079]
1930–40	0.0191 (0.0037)	−0.0048 (0.0014)	−0.0433 (0.0079)	0.0307 (0.0095)	0.71 [0.0041]
1940–50	0.0261 (0.0055)	−0.0135 (0.0022)	−0.0433 (0.0079)	0.0307 (0.0095)	0.82 [0.0065]
1950–60	0.0438 (0.0086)	−0.0205 (0.0031)	−0.0433 (0.0079)	0.0307 (0.0095)	0.70 [0.0091]
1960–70	0.0435 (0.0083)	−0.0056 (0.0025)	−0.0433 (0.0079)	0.0307 (0.0095)	0.70 [0.0069]
1970–80	0.0240 (0.0091)	−0.0077 (0.0024)	−0.0433 (0.0079)	0.0307 (0.0095)	0.73 [0.0072]
1980–89	0.0163 (0.0061)	−0.0066 (0.0019)	−0.0433 (0.0079)	0.0307 (0.0095)	0.72 [0.0053]
Joint, 8 subperiods	0.0260 (0.0023)	individual coefficients	−0.0427 (0.0079)	0.0300 (0.0097)	—

Note: The likelihood-ratio statistic for a test of the equality of the income coefficients over the 8 subperiods is 17.1, with a *p*-value of 0.017 (from a χ^2 distribution with 7 degrees of freedom). The regressions use iterative, weighted least squares and take the form.

$$m_{it} = a_t + b_t \cdot \log(y_{i,t-T}) + c_{1t} \cdot \text{heat}_i + c_2 \cdot \pi_{i,t-T} + c_3 \cdot \pi_{i,t-T}^2 + c_{4t} \cdot \text{region}_i + c_{5t} \cdot S_{it},$$

where m_{it} is the net flow of migrants into state i between years $t - T$ and t, expressed as a ratio to the population at $t - T$; heat_i is heating-degree days; $\pi_{i,t-T}$ is population density (thousands of persons per square mile); region_i is a set of dummies for the four main census regions; and S_{it} is the structural variable described in the text. The estimates of a_t, c_{4t}, and c_{5t} are not shown. The data are discussed in Chapter 10. All samples have 48 observations. Standard errors are in parentheses.

roughly 2 percent per year. The combination of the results for migration with those for income convergence suggests that net migration rates would be highly persistent over time. The data confirm this idea: the correlation between the average migration rate for 1900–40 with that for 1940–89 is 0.70.

11.6 MIGRATION ACROSS JAPANESE PREFECTURES

Before we analyze migration across Japanese prefectures and implement Eq. (11.9) for Japan, we should mention that there is a substantial difference between the typical Japanese prefecture and the typical U.S. state in terms of area. The average

size of a Japanese prefecture is 6,394 square kilometers,[11] roughly half the size of Connecticut. The largest prefecture, Hokkaido, is 83,520 km², or roughly the size of South Carolina. The second largest prefecture, Iwate, has an area of 15,277 km², a bit larger than Connecticut and a bit smaller than New Jersey. In comparison, the average U.S. state has an area of 163,031 km², and the area of the largest state in the continental United States, Texas, is 691,030 km². California, with an area of 411,049 km², is slightly larger than all of Japan (377,682 km²).

The contrast in size means that Japanese prefectures resemble metropolitan areas more than states, so that daytime commuting across prefectures can be significant. Urban economists, such as Henderson (1988), think that people like to live in cities for two reasons. First, there are demand or consumption externalities. That is, cities provide amenities, such as theaters and museums, features that can be supplied only if there is a sufficient scale of demand. Second, there are production external-ities, which tend to generate high wages in big cities. An offsetting force is that people want to live away from crowded cities because they tend to be associated with crime, less friendly neighborhoods, and (in equilibrium) high land and housing prices (see Roback [1982]). Thus, the decision to migrate to a city involves a tradeoff. This tradeoff can be avoided if people live in a suburb and commute to the central city. People are especially willing to pay high commuting costs when densities in the central city are extremely high.

To deal with these issues empirically, we would like to have a measure of the density of the neighboring prefectures. Conceptually, we could construct such a measure by weighting the neighbors' densities by their distance in some way. In practice, however, we observe that there are two main areas in Japan that have an abnormally high population density, Tokyo and Osaka. In 1990, Tokyo's density was 5,470 people/km² and Osaka's was 4,674 people/km², compared to an average for the other prefectures of 624 people/km².[12] Hence, the problems mentioned above are likely to arise in these two regions only. We can confirm this idea by considering the ratio of daytime to nighttime population, a measure of the extent of commuting.[13] A ratio smaller than one indicates that there are people who live in that prefecture but work in another, and a ratio larger than one indicates the opposite. The ratio is close to one for all prefectures except for the ones around Tokyo and Osaka: Tokyo's ratio is 1.184 and Osaka's is 1.053. The ratios for the Tokyo region are 0.872 for Saitama, 0.876 for Chiba, and 0.910 for Kanagawa. For the Osaka region, the ratios are 0.955 for Hyogo, 0.871 for Nara, and 0.986 for Wakayama.[14]

---

[11]This figure excludes Hokkaido, which is about five times as large as any of the other prefectures. The average size including Hokkaido is 8,036 km², two-thirds the size of Connecticut.

[12]In comparison, the U.S. state with the largest density in 1990 was New Jersey with 390 people/km².

[13]The source of these data is the Statistics Bureau, Management and Coordination Agency.

[14]There seems to be some commuting across prefectures in the areas surrounding Kyoto and Aichi, but the magnitudes are much smaller: Aichi's ratio is 1.016 (and its neighboring prefecture, Gifu, has a ratio of 0.977) and Kyoto's is 1.011.

We constructed a variable called "neighbor's density" by assigning the prefectures of the Tokyo area (Tokyo and its immediate neighbors, Saitama, Chiba, and Kanagawa) and the Osaka area (Osaka and its immediate neighbors, Hyogo, Nara, and Wakayama) the average density of their immediate neighbors. For other prefectures, the variable equals its own population density. We expect to find a positive relation between migration and this neighbor variable and a negative relation between migration and own density. This relation would indicate that people do not like to live in dense areas (they have to pay the congestion costs), but like to be close to these areas (so that they get the benefits of a big city).

The functional form that we estimate is

$$m_{it} = a + b \cdot \log(y_{i,t-T}) + c_1 \theta_i + c_2 \pi_{i,t-T} + c_3 \pi_{i,t-T}^{ne} + v_{it}, \quad (11.10)$$

where v_{it} is an error term, and $\pi_{i,t-T}^{ne}$ is the population density of the surrounding prefectures. To calculate the amenity (weather) variable, we squared the difference between the maximum and average temperatures, added the square of the difference between the minimum and average temperatures, and then took the square root. Hence, this variable measures extreme temperature. A variable similar to the one used for the United States (heating-degree days) was unavailable. We experimented with other weather variables, such as maximum and minimum temperatures and average snowfall over the year. These alternative variables did not fit as well.

Figure 11.11 shows the relation between the average annual migration rate for 1955–87 and the log of income per capita in 1955. The clear positive association (simple correlation of 0.58) suggests that net migration reacts positively to income differentials. An interesting point is that the three outliers at the top of the figure are Chiba, Saitama, and Kanagawa, the prefectures surrounding Tokyo.

Table 11.5 shows the results of estimating migration equations of the form of Eq. (11.10). The first row refers to the average migration rate for the whole period, 1955–90. The coefficient on the log of initial income per capita is 0.0126 (0.0061). As expected, net migration is negatively associated with own density (−0.0049 [0.0022]) and positively associated with neighbor's density (0.0190 [0.0034]). The extreme temperature variable is insignificant.

The next seven rows in Table 11.5 show results for the five-year subperiods beginning with 1955–60. The estimated coefficient on initial income is significantly positive for all subperiods, except for 1975–80, for which the coefficient is positive, but insignificant. The joint estimate is 0.0188 (0.0019), which implies that, other things equal, a 10 percent increase in a prefecture's per capita income raises net in-migration by enough to raise that prefecture's rate of population growth by 0.19 percentage points per year. This result is close to that found for the U.S. states. A test of the stability of the income coefficients over time yields a likelihood-ratio statistic of 18.0, with a p-value of 0.006.

The own-density variable is significantly negative, except for the first subperiod, and the neighbors' density variable is positive for all subperiods (significantly so for four of the seven subperiods). The extreme weather variable is negative, but only marginally significant. Thus, weather does not seem to play an important role in the process of internal migration in Japan.
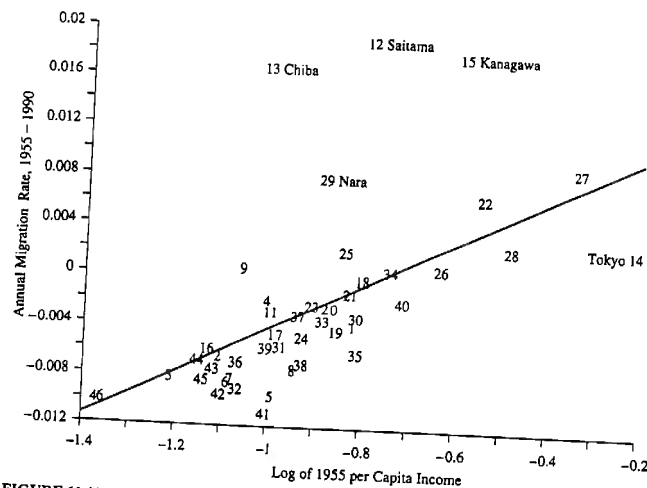
FIGURE 11.11

Migration and initial prefectural income, 1955–1990. The average net migration rate for 47 Japanese prefectures from 1955 to 1990, shown on the vertical axis, is positively related to the log of 1955 per capita income, shown on the horizontal axis. The three prefectures surrounding Tokyo—Chiba, Saitama, and Kanagawa—had substantially higher net migration rates than the values predicted by their initial levels of income.

To summarize, some main findings are that the rate of net in-migration to a prefecture is negatively related to own density and positively related to the density of neighbors. Holding other things constant, migration is positively associated with initial per capita income. A notable result is the similarity of the coefficients on income for the United States and Japan, 0.026 from the joint estimation for the U.S. states and 0.019 from the joint estimation for Japanese prefectures.

Recall that differences in per capita income tend to dissipate at a slow rate, something like 2.5 to 3 percent per year for the Japanese prefectures. Putting this result together with those for migration, the implication is that net migration rates would be highly persistent over time. The data confirm this idea: the correlation between the average migration rate for 1955–70 with that for 1970–90 is 0.60.

11.7 MIGRATION ACROSS EUROPEAN REGIONS

We now estimate the sensitivity of the net migration rate to income across the regions of the five large European countries: Germany, the United Kingdom, Italy, France, and Spain. The dependent variable is the average net migration rate for each of the

TABLE 11.5
Regressions for net migration into Japanese prefectures, 1955–90

Period	Log of per capita income	Extreme temperature	Own population density	Neighbors' population density	$R^2[\hat{\sigma}]$
1955–90	0.0126 (0.0061)	0.00014 (0.00062)	−0.0049 (0.0022)	0.0190 (0.0034)	0.62 [0.0061]
1955–60	0.0216 (0.0036)	−0.00014 (0.00012)	0.0060 (0.0013)	0.0025 (0.0019)	0.85 [0.0038]
1960–65	0.0317 (0.0058)	−0.00014 (0.00012)	−0.0019 (0.0020)	0.0147 (0.0031)	0.74 [0.0071]
1965–70	0.0344 (0.0070)	−0.00014 (0.00012)	−0.0065 (0.0017)	0.0142 (0.0025)	0.71 [0.0066]
1970–75	0.0194 (0.0060)	−0.00014 (0.00012)	−0.0064 (0.0015)	0.0114 (0.0023)	0.53 [0.0070]
1975–80	0.0060 (0.0067)	−0.00014 (0.00012)	−0.0037 (0.0011)	0.0052 (0.0014)	0.32 [0.0043]
1980–85	0.0101 (0.0044)	−0.00014 (0.00012)	−0.0023 (0.0006)	0.0037 (0.0086)	0.39 [0.0030]
1985–90	0.0148 (0.0040)	−0.00014 (0.00012)	−0.0026 (0.0006)	0.0046 (0.0084)	0.56 [0.0029]
Joint, 7 subperiods	0.0188 (0.0019)	−0.00040 (0.00015)	individual coefficients	individual coefficients	—

Note: The likelihood-ratio statistic for the hypothesis that the income coefficients are the same is 18.0, with a p-value of 0.006. The regressions use iterative, weighted least squares to estimate equations of the form,

$$m_{it} = a_t + b \cdot \log(y_{i,t-T}) + c_1 \cdot \text{temp}_i + c_{2t} \cdot \pi_{i,t=T} + c_{3t} \cdot \pi_{i,t-T}^{ne} + c_{4t} \cdot \text{district}_i + c_{5t} \cdot S_{it},$$

where m_{it} is the net flow of migrants into prefecture i between years $t - T$ and t, expressed as a ratio to the population at time $t - T$; temp_i is a measure of extreme temperature, calculated as deviations of maximum and minimum temperatures from the average temperature; $\pi_{i,t-T}$ is population density (thousands of persons per square kilometer); $\pi_{i,t-T}^{ne}$ is the population density of the neighboring prefectures (see the text); district_i is a set of dummy variables for the district; and S_{it} is the structural variable described in the text. All samples have 47 observations. (See the note to Table 11.4 for additional information.)

four decades starting in 1950 (see Chapter 10 for a discussion of these data). We are missing observations for the United Kingdom in the 1950s and 1980s and for France in the 1980s.

We estimate a system of regressions similar to those used for the United States and Japan. The explanatory variables are the logarithm of per capita GDP or income at the beginning of the decade, population density at the beginning of the decade, sectoral variables (shares in employment or GDP of agriculture and industry at the start of each decade), a temperature variable, and country dummies. We estimate a system of equations for the five countries, with the density and temperature variables restricted to have the same coefficients over time and across countries, but with the coefficients of the other variables allowed to vary over time and across countries.

TABLE 11.6
Regressions for net migration into European regions, 1950–90, coefficients on the log of per capita GDP

	1950s	1960s	1970s	1980s	Total
Germany	0.0311 (0.0121)	0.0074 (0.0088)	0.0040 (0.0038)	0.0024 (0.0086)	0.0076 (0.0014)
United Kingdom	—	0.0049 (0.0011)	−0.0069 (0.0013)	—	−0.0041 (0.0023)
Italy	0.0182 (0.0041)	0.0208 (0.0027)	0.0089 (0.0020)	0.0309 (0.0106)	0.0117 (0.0018)
France	0.0090 (0.0056)	−0.0008 (0.0095)	0.0097 (0.0041)	—	0.0100 (0.0036)
Spain	0.0126 (0.0068)	0.0135 (0.0112)	0.0117 (0.0063)	0.0031 (0.0070)	0.0034 (0.0021)
Overall	0.0107 (0.0038)	0.0072 (0.0040)	0.0046 (0.0024)	0.0141 (0.0070)	0.0064 (0.0021)

Note: The regressions take the form,

$$m_{ijt} = a_{jt} + b_{jt} \cdot \log(y_{ij,t-T}) + c_1 \cdot \text{temp}_{ij} + c_2 \cdot \pi_{ij,t-T}$$
$$+ c_3 \cdot (\text{country dummy}) + c_{4jt} \cdot \text{AG}_{ij,t-T} + c_{5jt} \cdot \text{IN}_{ij,t-T},$$

where m_{ijt} is the net flow of migrants into region i of country j between years $t - T$ and t, expressed as a ratio to the population at time $t - T$; temp_{ij} is the average maximum temperature; $\pi_{ij,t-T}$ is population density (thousands of persons per square kilometer); $\text{AG}_{ij,t-T}$ is the share of employment or GDP (for the 1980s) in agriculture; and $\text{IN}_{ij,t-T}$ is the share in industry. All estimation is by the iterative, seemingly unrelated procedure. The table reports only the estimates of the coefficients b_{jt}. The numbers in the first five rows and first four columns apply when each country has a different coefficient for each period. The last column restricts the coefficients to be the same over time for each country. The last row restricts the coefficients to be the same across countries for each decade. The number in the intersection of the last row and column applies when all countries and time periods have a single coefficient.

Table 11.6 reports the estimated coefficients on the log of initial per capita GDP or income. The first column contains the estimates for the 1950s, the second for the 1960s, and so on. The last column restricts the coefficients to be the same over the decades. The first row is for Germany, the second for the United Kingdom, the third for Italy, the fourth for France, and the fifth for Spain. The last row restricts the coefficients to be the same for the five countries.

In contrast with the results for the United States and Japan, the coefficients on the log of per capita GDP or income are not precisely estimated for the European countries. For Germany, the estimated coefficient for the 1950s is positive and significant, 0.031 (0.012), whereas those for the other three decades are insignificant. The estimated income coefficients for Italy are significantly positive, but many of those for the United Kingdom, France, and Spain are insignificant.

If we restrict the coefficients to be the same over time, but allow them to vary across countries, then the estimated values are 0.0076 (0.0014) for Germany, −0.0041 (0.0023) for the United Kingdom, 0.0117 (0.0018) for Italy, 0.0100 (0.0036) for France, and 0.0034 (0.0021) for Spain. If we restrict the coefficients to be the same across countries, but allow them to vary over time, then the estimated

values are 0.0107 (0.0038) for the 1950s, 0.0072 (0.0040) for the 1960s, 0.0046 (0.0024) for the 1970s, and 0.0141 (0.0070) for the 1980s. Finally, if we restrict the coefficients to be the same across countries *and* over time, then we get 0.0064 (0.0021). Although this estimate is significantly positive, the size of the coefficient is much smaller than the comparable values for the United States (0.026) and Japan (0.019). The main finding therefore is that the migration rate for European regions is positively related to per capita GDP or income, but the magnitude of the relation is weak, and the coefficients cannot be estimated with great precision.

11.8 MIGRATION AND CONVERGENCE

We found in Chapter 9 that the migration of workers with low human capital from poor economies to rich economies tends to speed up the convergence of per capita income and product. The convergence coefficients estimated in growth regressions would include this effect from migration. In this section, we attempt to estimate the effect of migration on convergence by including the net migration rate as an explanatory variable in the growth regressions. If migration is an important source of convergence—and if we can treat the migration rate as exogenous with respect to the error term in the growth equation—then the estimated convergence coefficient, β, should become smaller when migration is held constant.

We enter the contemporaneous net migration rate in growth regressions in Table 11.7. The first row reports the estimated speed of convergence, β, for the U.S. states. The sample period, 1920–90, is divided into seven ten-year subperiods. The regression includes period-specific coefficients for constant terms, dummies for the four major census regions, and the structural variable discussed before. The coefficient on the log of initial per capita income is constrained to be the same for each subperiod. This setup parallels the joint estimation shown in Table 11.1, column 3, except for the elimination of the two early subperiods.

Column 1 of the table reports the estimate of β when the migration rate is not included in the regressions. The speed of convergence is 0.0196 (0.0025), close to the familiar 2 percent per year. Column 2 adds the net migration rate as a regressor. (The coefficient on this variable is constrained to be the same for each subperiod.) The estimated coefficient on the migration rate is positive and significant, 0.093 (0.030), and the estimate of β, 0.0231 (0.0028), is actually somewhat higher than that shown in column 1. Thus, contrary to expectations, the estimate of β does not diminish when the net migration rate is held constant.

The results are likely influenced by the endogeneity of the net migration rate. Specifically, states with more favorable growth prospects (due to factors not held constant by the included explanatory variables) are likely to have higher per capita growth rates *and* higher net migration rates. We attempt to isolate exogenous shifts in migration by using as instruments the explanatory variables used to explain the net migration rate in Table 11.4. These variables include population density and the log of heating-degree days. (The assumption here is that some of these determinants of migration do not enter directly into the growth equation.) The results, contained in column 3 of Table 11.7, show an insignificant coefficient on the migration rate, −0.006 (0.048), and an estimated β coefficient, 0.0174 (0.0033), that is slightly

TABLE 11.7
Migration and convergence

	(1) Migration excluded	(2) Migration included (OLS)		(3) Migration included (IV)	
	β	β	Migration	β	Migration
United States 1920–90	0.0196 (0.0025)	0.0231 (0.0028)	0.0931 (0.0305)	0.0174 (0.0033)	−0.006 (0.048)
Japan 1955–90	0.0312 (0.0040)	0.0340 (0.0044)	0.0907 (0.0041)	0.0311 (0.0042)	−0.108 (0.112)
Germany 1950–90	0.0243 (0.0088)	0.0240 (0.0091)	−0.014 (0.235)	0.0181 (0.0093)	−0.542 (0.429)
United Kingdom 1960–80*	0.0176 (0.0132)	0.0220 (0.0203)	0.116 (0.395)	0.0261 (0.0267)	0.222 (0.570)
Italy 1950–90	0.0206 (0.0058)	0.0244 (0.0070)	0.166 (0.156)	0.0180 (0.0098)	−0.121 (0.370)
France 1950–80**	0.0224 (0.0265)	0.0172 (0.0063)	−0.038 (0.126)	0.0177 (0.0065)	−0.084 (0.178)
Spain 1950–90	0.0245 (0.0102)	0.0295 (0.0096)	−0.124 (0.102)	0.0268 (0.0119)	−0.068 (0.203)

Note: The regressions for the growth rates of per capita income or GDP are analogous to the joint estimations shown in Table 11.1, column 3 for the U.S. states; Table 11.2, column 3 for the Japanese prefectures; and Table 11.3, column 2 for the European regions (except that the five large European countries are treated separately here). The β coefficients refer to the log of initial per capita income or GDP, and the migration coefficients refer to the net migration rate. In column 1, the migration rate is not included as a regressor. In column 2, the migration rate is added, and the estimation is by OLS. In column 3, instrumental estimation is used. The instruments are the regressors included in the migration equations, as reported in Table 11.4 for the United States, Table 11.5 for Japan, and Table 11.6 for Europe.

*Two subperiods.

**Three subperiods.

lower than that in column 1. These results suggest that migration does not account for a large part of β convergence for the U.S. states.

The second row of Table 11.7 applies the same procedure to Japan. The first column reports the joint estimate of β over seven five-year periods when the migration rate is excluded as a regressor. The estimate of β, 0.0312 (0.0040), is the same as that in column 3 of Table 11.2. When the migration rate is added in column 2 of Table 11.7, the estimated coefficient on migration is positive and similar to that found for the United States, 0.0907 (0.0041), and the estimate of β increases to 0.0340 (0.0044). In column 3, which includes instruments for migration, the estimated coefficient on migration is insignificant, −0.11 (0.11), and the estimate of β, 0.0311 (0.0042), is essentially the same as that in column 1. Hence, as for the U.S. states, migration does not appear to be a major element in β convergence for the Japanese prefectures.

The last five rows of Table 11.7 apply an analogous procedure to the five large European countries. The main findings are similar to those for the United States and Japan in that the estimated β coefficients do not change a great deal when migration

rates are held constant. One surprising result here is that the net migration rates are insignificant in the OLS regressions for the European regions, whereas the usual endogeneity story suggests positive coefficients. It may be that the regional net migration rates are not well measured for the European countries, a possibility that would also account for the difficulties in the estimated migration equations in these cases.

A second prediction from the migration theory in Chapter 9 is that economies with higher sensitivity of net migration to per capita income will have higher convergence coefficients, β. To check this possibility, we plot in Fig. 11.12 the estimated β coefficients against the estimated coefficients of the log of per capita GDP or income from the migration equations. The figure has seven data points, corresponding to the United States, Japan, Germany, the United Kingdom, Italy, France, and Spain. The figure shows a weak positive relation between the two coefficients; the correlation is 0.27.[15] The imprecision with which the coefficients in the migration equations are
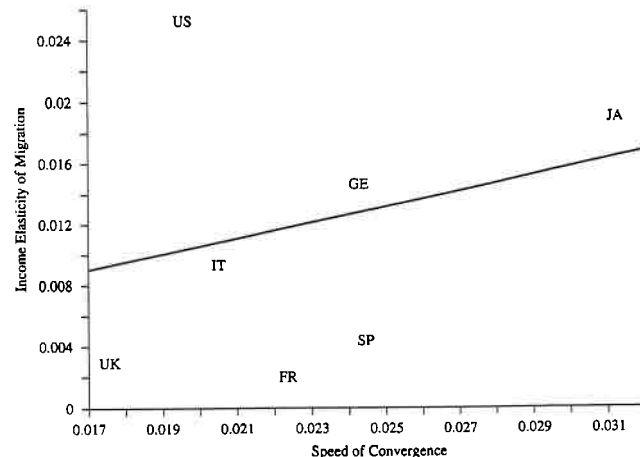

FIGURE 11.12
Income coefficient of migration and speed of convergence. The vertical axis shows the estimated coefficient on the log of per capita income or GDP from migration regressions. The horizontal axis has the estimated β convergence coefficient from growth regressions. The seven data points—for the United States, Japan, Germany, the United Kingdom, Italy, France, and Spain—exhibit a positive relation, as predicted by the theory of migration and growth.

[15]The β coefficients for France and the United Kingdom are those estimated over the same subperiods for which the migration data are available. The β coefficient estimated over the full sample is lower for France and higher for the United Kingdom. If we use these alternative estimates of β, then the correlation with the coefficient from the migration equations is slightly higher, 0.32.

estimated for the European countries suggests that this relation should be interpreted with caution. See Braun (1993) for further discussion of this approach.

11.9 CONCLUSIONS

We studied the behavior of the U.S. states since 1880, the prefectures of Japan since 1930, and the regions of eight European countries since 1950. The results indicate that absolute β convergence is the norm for these regional economies. That is, poor regions of these countries tend to grow faster per capita than rich ones. The convergence is absolute because it applies when no explanatory variable other than the initial level of per capita product or income is held constant.

We can interpret the results as consistent with the neoclassical growth model described in Chapters 1 and 2 if regions within a country have roughly similar tastes, technologies, and political institutions. This relative homogeneity generates similar steady-state positions. The observed convergence effect is, however, also consistent with the models of technological diffusion described in Chapter 8.

One surprising result is the similarity of the speed of β convergence across data sets. The estimates of β are around 2–3 percent per year in the various contexts. This slow speed of convergence implies that it takes 25–35 years to eliminate one-half of an initial gap in per capita incomes. This behavior deviates from the quantitative predictions of the neoclassical growth model if the capital share is close to one-third. The empirical evidence is, however, consistent with the theory if the capital share is around three-quarters.

The analysis of migration indicates that the rate of net migration tends to respond positively to the initial level of per capita product or income, once a set of other explanatory variables is held constant. This relation is clear for the U.S. states and the Japanese prefectures, but is weaker for the regions of five large European countries. We also check whether the presence of β convergence in the regional data can be explained by the behavior of net migration. The evidence here is not definitive, but suggests that migration plays only a minor role in the convergence story.