
18 Immigration and the Receiving Economy

Rachel M. Friedberg and Jennifer Hunt

THE PUBLIC DEBATE OVER immigration policy in the United States has become quite heated in recent years. The passage of Proposition 187 in California in 1994, making illegal aliens ineligible for public health and education services, and the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, which curtailed immigrants' eligibility for public assistance, may herald the beginning of a new period of tighter restrictions on immigration and immigrant rights in the United States.

In assessing where the U.S. government should stand on issues related to immigration policy, it is important to have a firm grasp of the facts concerning immigration's effect on the U.S. political system, society, and economy and to sort out the facts from the myths. This chapter is about the economic dimension of immigration's impact on the United States.

Those who favor tight restrictions on immigration often argue that immigration has strong detrimental effects on the U.S. economy. They assert that immigrants have an adverse impact on the labor market outcomes of the native-born population, and that through the competition they create in the labor market, the job opportunities of the U.S. population suffer. Groups lobbying against immigration argue that we must guard against these detrimental consequences of immigration in our country. They assert that not only do immigrants take jobs away from American workers, but immigrants also drag down levels of pay, toward the standard of living of the countries from which they come.

Some supporters of relatively liberal entry laws contend that, despite these potential economic costs, it is our moral obligation as a nation to share the political freedom and economic opportunity we enjoy with others who wish to live here. Yet other advocates of more open immigration argue that the opponents of immigration have the economic evidence all wrong. They assert that immigrants do not compete with native-born workers but rather fill occupational slots that native-

born workers would not be willing to take. They add that immigrants provide these services at a low cost to American employers and consumers. Moreover, they argue, immigrants revitalize our economy with their initiative and entrepreneurial spirit and contribute to economic growth.

This chapter outlines the major questions addressed in the economics literature concerning the impact of immigration on the receiving economy. The emphasis is on the effects of immigration on the labor market outcomes of the receiving population, including the impact on their wages, employment rates, unemployment rates, income growth, and migration patterns. The following section discusses the theory used in analyzing these questions, while the subsequent section provides a critical overview of the existing empirical evidence. Immigration has been an area of very active research in economics in recent years: most of the studies surveyed here have come out within the last five years.

To preview the conclusion of this chapter, to the extent that the literature in this area can be said to be in consensus, its conclusion is that immigration does not have significant adverse effects on the receiving economy. Immigrants are not found to cause a drop in the employment rates of the U.S. labor force, and although some studies find economically significant negative effects of immigrants on native wages, these magnitudes are very likely overestimates.

ECONOMIC THEORY FOR THE STUDY OF IMMIGRATION

Theory

Economics does not have a body of theory developed specifically for the purpose of studying immigration. Rather, immigration is examined in the context of the field of economics, which studies the particular aspect of immigration under consideration, using the tools and methodology standard

to that field. Analyzing the effect of immigration on tax revenue and public assistance expenditures falls under the rubric of public finance; the consequences of immigration for economic growth call for the framework of macroeconomics; and the impact of immigrants on the employment opportunities of the native-born is a question for labor economics. Advances or changes in the approaches taken to studying immigration therefore follow changes in the field more generally. Almost no new theory has been developed specifically for obtaining answers to questions of immigration.

Models

Wages and Employment The fundamental model used to evaluate the labor market impact of immigration is that of supply and demand in the market for labor services. The interaction of sellers (workers) and buyers (employers) leads to a certain volume of transactions (the level of employment) at a certain price (the going wage rate). Labor economics studies the determination of the equilibrium price and quantity transacted in this market.

The values of those two variables depend on supply-and-demand conditions and on their interaction. A large supply of a good tends to depress its price and raise the quantity transacted, while scarcity on the supply side tends to raise the price of the good and lower the quantity transacted. High demand for a good tends to raise its price and raise the quantity traded. Conversely, low demand tends to lower the price and lower the quantity traded.

The most salient way in which immigration fits into this model is that immigration represents an increase in the number of workers interested in working at any given level of the wage rate. The result is an outward shift of the labor supply curve. According to the basic model of supply and demand, if there is no accompanying change in demand, this increase in supply will lead to a fall in the wage and a rise in the total number of people employed. (This is a movement down the labor demand curve, which indicates for any given wage how many workers firms will want to employ.) However, the rise in employment will not be as large as the immigration. Because the wage has fallen, some people will no longer desire to work—these may be immigrants or natives. If some are natives, the native employment rate (the ratio of employed people to population) will fall. From the employers' point of view, the wage has

not fallen enough to warrant accommodating the full increase in labor supply. Total output is higher: because inputs have become cheaper, optimal output levels for firms rise. Looked at this way, the increase in employment may be said to be due to a "scale effect."

This analysis may be made richer by consideration of other factors of production—namely, inputs used in the production process such as natural resources, capital, and different types of workers. The analysis so far implicitly assumes that all production takes place with one type of worker, and that workers, both native and immigrant, are identical. If natives and immigrants are identical, they may be called perfect substitutes. There may instead be two types of workers in production, skilled and unskilled, which are imperfect substitutes in production. A firm could produce the same output by reducing the number of one type of worker and increasing the number of the other, but it will not replace one for one, as an additional worker has a different effect on production depending on his or her type.

In many situations, immigrants are unskilled, so we can model immigration as an expansion in the supply of unskilled workers (assuming that unskilled immigrants and unskilled natives are identical, or perfect substitutes). As in the simpler case, this expansion will lead to a fall in the wages of the unskilled, and the desired output level of firms will rise, since inputs have become cheaper (scale effect). This rise in output is likely to increase demand for skilled workers, tending to raise their wage. However, unskilled workers are now a more attractive option in the production process relative to the skilled, since their wages have fallen, and thus firms will shift the mix of workers away from skilled workers and toward unskilled workers (the "substitution effect"). This effect will tend to decrease demand for skilled workers, thus decreasing their wage. Hence, the net effect of unskilled immigration on the wages of skilled natives is ambiguous; it depends on the relative sizes of the scale and substitution effects. In general, immigrants reduce the wages of workers for whom they are perfect substitutes and have an ambiguous effect on the wages of those for whom they are imperfect substitutes.

Immigration increases the return to factors with which immigrants are complements. Two factors are called complements when they must be used together in production—for example, scalpels and surgeons. It is thought, more generally, that capital equipment and skilled workers may be comple-

ments: if a firm reduces the amount of capital it uses, it cannot keep output constant by increasing the number of skilled workers. Without capital to work with, skilled workers cannot contribute to production.

Consider the case of immigration of people who are perfect substitutes for skilled natives. Using reasoning similar to that of the previous case, we see that this will cause a fall in the wages of skilled workers. This wage reduction will cause a scale effect, inducing firms to use more of all inputs and tending to raise wages of unskilled workers and the return to capital. As skilled workers become relatively cheaper, firms will substitute skilled workers for unskilled ones, leaving the net effect on the wages of the unskilled ambiguous. However, since capital is complementary to skilled labor, there is no such substitution of skilled labor for capital: rather, the increased use of skilled labor leads to an increase in the use of capital. The return to capital therefore unambiguously rises. In general, immigration raises the return to factors with which it is a complement.

Popular fears of the effects of immigrants on wages often center on the idea that immigrants bid down wages by being prepared to work for less than natives. (Economists express this idea by saying that immigrants may have lower reservation wages than natives.) This is particularly likely in the case of illegal immigrants, whose options are more limited than those of comparable natives. If this is true, it means that immigrants have a steeper labor supply curve than natives: at low wages few natives but more immigrants will want to participate in the labor force, while at high wages most of both groups will want to participate. So when immigrants arrive, they increase labor supply more at the low end, thus decreasing wages further than an inflow of people identical to natives would have done.

The theory outlined so far assumes that the receiving economy does not trade with the rest of the world. The realization that countries trade with one another leads to models that have strong predictions about wages and prices throughout the world: the most simple trade model predicts that wages, prices, and the return to capital will be equalized across the world (so-called factor price equalization). The factor price equalization prediction comes from the following reasoning. With free trade, the relevant market for a good is the world market, and the price will be set in this market. Just as there would be a single price for a good in a closed economy with many firms, with

free trade there will be a single world price for a good, determined by world supply and world demand by firms and consumers in many countries. The borders are irrelevant. It is assumed that no one country produces enough of a good to be able to alter the world price through changes in production.

If firms in all countries use the same technology (and this is an important and contentious assumption), labor demand for each firm (how much labor is demanded at any given wage level), which is determined by technology and the price of output, will be the same in all countries producing the same set of goods. This will lead to wages being equal in all countries: if one country has less labor, this would appear in our closed economy analysis to lead to the wage being higher in that country. However, this higher wage will induce firms in that country to increase production of goods using more capital and to reduce production of goods using labor intensively, thus decreasing labor demand and leading to the wage falling to the world level. Thus, a country's endowment of labor and capital will determine which goods it exports and which it imports. This luxury of adjusting the product mix with no change in goods prices was not available to the closed economy.

According to this analysis, immigration has no effect on wages in receiving or sending countries. An influx of labor into a country will cause it to produce more labor-intensive goods and less capital-intensive goods. Through this adjustment, the wage will return to its old level: the adjustment may be thought of as an export of immigrants in the form of higher exports (or lower imports) of labor-intensive goods. A similar analysis may be applied to emigration, and indeed, the whole analysis holds for prices and movements of capital.

Notice that factor price equalization, which predicts that immigration will not affect wages, also implies that there is no (economic) incentive for migration. Adding the real-world feature of tariffs to the model, however, can generate migration from poor to rich countries. Rich countries often impose tariffs on the import of labor-intensive goods (raising their domestic price above the world price). When this tariff induces domestic firms to produce more of the labor-intensive good, labor demand and hence wages both rise. This rise will occur (although it will not be beneficial to the country as a whole) if capital and labor cannot cross borders. If immigration is possible, however, people will seek to move to countries with such tariffs, lowering the wage. (Since the domestic

price of the protected good can fall, the analysis in this situation is similar to the closed economy analysis.) Immigration will continue until the wage returns to the world level (by which time labor will be so abundant that the country will specialize in the production of that good).

The predictions of this most simple trade model are thought to be unrealistic, but even more complex versions of the model predict that small numbers of immigrants will have no effect on the wage. The initial analysis relied on the assumption that countries have similar enough factor endowments to produce the same goods. If there is an array of products with technologies requiring different intensities of labor and capital inputs, the mix of products a country produces will be determined by its endowment of factors. Small inflows of immigrants will lead to adjustment in production and trade of the existing goods, and the wage will be unaffected. However, a large enough inflow will lead to a fall in the wage and to a change by the country in its product mix toward more labor-intensive products.

Another assumption of the simple model that appears unrealistic is that all countries produce a given good with the same technology. If some countries have a technological advantage in the production of, say, labor-intensive goods, wages will be higher, inducing immigration to those countries. As in the simple trade model, this immigration will not affect the wage until the mix of products changes to be more labor-intensive. The advantage of this model is in proposing an economic incentive for migration (see Markusen 1983).

If the country receiving immigrants is a large one, the increase in output of labor-intensive goods spurred by immigration will be large enough to reduce the world price of these goods. This may reduce the wage, even under conditions where factor price equalization would hold for a small country.

Theoretical models thus have predictions ranging from a definite fall in the wage and employment rate (the closed economy case) to a wage fall after a certain threshold of immigration is passed, to no fall in wages (open economy models). Empirical work is therefore needed to resolve the question of whether immigration has any effect, and if so, how much.

Unemployment For theoretical purposes, a person is said to be unemployed if he or she would like to work at the wage offered to those with a

job but cannot find a job. For empirical or, more simply, measurement purposes, emphasis is placed on the search process: only jobless persons actively searching for a job are unemployed. Others are considered out of the labor force, even if they would accept a job were it offered to them. Studies often measure the impact of immigration on the unemployment rate, which is defined as the ratio of the number of unemployed to the number in the labor force. An unemployed person is considered to be in the labor force, even though he or she does not have a job, so the labor force is the sum of employed and unemployed individuals. The labor-force participation rate, another variable of interest in the immigration context, is thus the ratio of those in the labor force to the population.

Simple economic models of labor supply and demand do not predict the existence of unemployment but rather predict that wages will adjust so that all who want to work at the going wage are able to do so. Many models analyzing the effect of immigration on wages hence have implications about employment, but not about unemployment. The closed economy model predicts that the fall in the wage allows employment to rise, but by less than the amount of the immigration, implying a fall in the labor-force participation rate. The simple open economy model does not incorporate a labor-force participation decision and results in trade adjustments, permitting employment to rise by the amount of the immigration.

The simplest explanation of the existence of unemployment predicts "frictional" unemployment. If an individual does not have a job, and wishes to find one, he or she will not have perfect information about what is available, and furthermore, applications can introduce delays, so that the job search will take some time. Thus, at any time a certain amount of unemployment will exist due to workers taking some time to discover and choose among possibilities and going through selection procedures established by firms. There are several reasons why individuals without a job may be searching for one: one example is entrants to the labor market, either first-time entrants who just finished their studies or people who temporarily left the labor force and now wish to work again. Another important category of unemployed workers are those who might have lost their previous job owing to "sectoral shifts." Over time the fortunes of sectors or industries wax and wane, possibly owing to technological innovation, for example. In recent years, the computer industry has grown and taken on new workers, while the type-

writer industry has shrunk and laid off workers. Owing to the changing fortunes of industries, there will always be a certain amount of churning in the labor market. The changing fortunes of firms within an industry will have the same effect.

Both open and closed economy models predict that workers and capital will shift toward labor-intensive sectors in response to immigration, and even that new sectors will be established as the product mix changes. In the short run, this shift is likely to cause some frictional unemployment as workers look for new jobs and capital relocates. The immigrants themselves will, of course, suffer more unemployment than natives in the short run, since all of them wishing to work need to search for jobs, while only some of the natives will need to change sectors.

Economists have other explanations for unemployment, explanations that in turn have implications for the impact of immigration. Theory predicts, for example, that the minimum wage should cause unemployment. Minimum-wage legislation arises in response to a perception that the market wages for some groups are too low and that a higher wage should be legislated. If the minimum wage for a group is set above the market wage, the amount of labor demanded will fall to below what it would have been at the market wage. By contrast, the labor supplied will rise with the wage. Employment will obviously equal labor demanded, and hence a group of people who want to work at the offered wage will not be able to and are thus unemployed. There may be different minimum wages affecting different groups: some states in the United States have one general minimum wage, for example, and a subminimum wage for younger workers. In some countries and in some periods, minimum wages by occupation have been set.

The effect of union bargaining on unemployment is conceptually similar to that of the minimum wage. If unions bargain a wage above the market wage, employment will fall and unemployment will occur. Unions should cause unemployment only if they are active in all sectors of the economy. If there are nonunionized sectors, people displaced from union sectors may seek jobs in these sectors, pushing down the wage there. Some may prefer to keep trying for a better-paying union job, however, causing "wait" unemployment. Unions may negotiate wages by industry or occupation, or they may take age, experience, or tenure into account in setting the wage.

If there is immigration of people who are perfect substitutes for workers facing a wage fixed

above the market rate, either a minimum wage or a union contract wage, there can be no downward adjustment of the wage. The amount of labor supplied will increase, and since the wage does not respond, unemployment will increase. It is again likely that in the short run the unemployment will disproportionately affect immigrants.

A more complex set of ideas about unemployment comes under the heading of "efficiency wages." These seek to explain why employers might voluntarily pay a wage above the market wage, a practice that would cause unemployment if practiced by all firms. (If not, workers should move to other firms, as in the union case.) The original motivation came from the development economics literature, which postulated that paying workers more might allow them to feed themselves better and make themselves more productive. In a developed country context, it is surmised that if firms cannot monitor their employees' output perfectly or cheaply, a good strategy may be payment of a higher wage to motivate them to work better (Shapiro and Stiglitz 1984). Then workers have more incentive to work harder: if they lose their job, they will face a wage reduction. Since if all firms do this unemployment arises, an additional motivation is the fear of unemployment if shirking is discovered. A related motivation may be a desire of firms to avoid quits, which they may find costly. Yet another hypothesis, the "fair wage" hypothesis, is that firms choose to pay efficiency wages to their less skilled workers to avoid the possibility of them becoming disgruntled at the wage gaps between the more and less skilled workers in the firm. Finally, firms may choose to pay efficiency wages in order to discourage their workers from trying to form a union, which might be even more costly to firms than efficiency wages.

If immigration occurs in a situation of efficiency wages, the increase in the size of the labor force allows firms to lower the wage and raise employment. However, the lower wage must be accompanied by a rise in the unemployment rate to maintain workers' incentive not to shirk. How the rise in unemployment is distributed across natives and immigrants is an empirical question.

Several unemployment theories thus suggest that immigration will increase unemployment. Efficiency wages may lead to higher long-run unemployment. The minimum wage may also do so if it is indexed to the real wage, as in France; if it has a fixed nominal level, however, as in the United States, it becomes less binding over time. New legislation may take the increase in labor supply

into account. Likewise, unions may eventually take the increase in labor supply into account. An increase in frictional unemployment, however, should be only temporary.

Economic Growth The determination of growth in a country's average income per person, including the influence of immigration on that process, has been a subject of interest among macro and international economists rather than labor economists. Income may be thought of as the sum of wages and the return to capital, and income growth is likely to imply higher wages and returns to capital, although in principle it could come from growth in one or the other. The starting point of growth analysis is a model developed in the 1950s by Robert Solow (1956). A typical modified model assumes that production takes place using (physical) capital, labor, and human capital. Human capital refers to the knowledge or skill in which individuals may invest in order to make themselves more productive. It is assumed that individuals (labor) may migrate, bringing their human capital with them, but that physical capital cannot move internationally, and furthermore, that there is no trade between countries.

In this model, immigrants migrate to countries with high ratios of physical capital to labor, and hence high wage rates. The more capital a worker uses, the more productive he or she is, and hence the higher the wage he or she earns; the higher wage naturally attracts immigrants. The higher the ratio of physical to human capital, and hence the return to human capital, the more human capital the immigrants will bring with them. Analogous reasoning applies here: if the available skill in a country is matched with a lot of capital, it will be very productive and small increases in skill will bring significant productivity and hence wage increases. Thus, in this situation it will be particularly worthwhile for individuals with a lot of skill (human capital) to immigrate.

The key issue in the impact of immigration is whether immigrants bring enough human capital to offset their dilution of physical capital in the receiving economy. All else (and human capital in particular) being equal, immigration will reduce the productivity of workers, since the amount of capital per worker will fall and hence wages will fall. However, the increased human capital that might result from immigration is assumed to lead to an increase in the capital stock and makes all workers more productive. If immigrants have little human capital, their impact is akin to that of faster

population growth in slowing per capita growth. If immigrant human-capital levels are higher than natives' by a sufficient amount, growth will be speeded up. Models of this type, as well as extensions that show how much people in a country save, have the obvious limitation of ignoring trade and capital mobility, although some steps have been taken to rectify this (see Braun 1992).

A common assumption in economics is that of constant returns to scale: if all inputs to production are changed by some multiple, output will be changed by that multiple. Intriguing outcomes arise if the possibility of increasing returns to scale is considered. With increasing returns to scale, the doubling of all inputs, for example, will lead to more than doubling of output. Thus, immigration to a country will increase output more than proportionately, raising both the wage and the return to capital and causing growth in per capita income (see Brezis and Krugman 1996).

METHODOLOGIES

Most of the recent research activity in the economics of immigration has not focused on improving existing theories of the economy but rather has been directed toward improving our ability to measure empirically the effects of immigration. The major developments in the field have been in the area of methodological approaches to empirical research. Researchers aim to measure the changes in economic variables that can be attributed to immigration. An underlying presumption of the research in economics is that the mechanisms at work are general ones, and that, with certain caveats, the conclusions of these studies can be applied to other labor market settings as well (other countries, time periods, and so on). This section reviews and critiques the empirical approaches used in the recent literature in the field and surveys the findings of that work.

Empirical economics involves the application of econometrics—tools of statistical analysis—to data on the exogenous and endogenous variables of interest, in order to generate predictions about the state of the endogenous variables. The question posed is usually of the form "Holding all other factors constant, what happens to variable Y when there is a change in variable X ?" In this case, the questions are: What happens to various measures of economic conditions when there is a change in the level of immigration? The quality of an empirical study clearly depends on the quality

of the model, the quality of the data, and the degree to which the assumptions under which the econometric technique yields unbiased results are satisfied.

A fundamental problem faced by social scientists is that, unlike researchers in the hard sciences, they cannot generally perform controlled experiments. The variable of interest is therefore not the only factor changing, so that it is difficult to isolate its effect. In the absence of controlled experiments, the next best alternative is to observe a natural experiment. Later in this section, we review some analyses of natural experiments relevant to the study of immigration. Most empirical economic research, however, has neither artificial nor natural experiments to use and must turn to other methods. In the area of immigration, the following techniques have been applied.

Exploiting the Geographic Dimension

In order to conduct an empirical investigation of the hypothesis that immigrant inflows have an impact on local economic conditions, it is necessary to collect data that display variation in the variables of interest. That variation can take place across any of several dimensions, including time or space. (Other dimensions could include different skill levels in the workforce, different industrial sectors of the economy, and so on.) This section describes the research that is more heavily based on spatial rather than temporal variation.

The Single Cross-sectional Approach Studies of immigration that emphasize geographic variation often use as their unit of observation the region, state, or SMSA (Standard Metropolitan Statistical Area), which encompasses the entire metropolitan area of a city. A cross-sectional data set is one in which observations on the variables of interest are made for many individual or geographic units at the same point in time. An example of such a data set is the Public Use Microdata Sample of the decennial U.S. census of population. Using this data set, in which the unit of observation is individual people, one can construct state or city-level measures of the fraction of the population that is foreign-born, the percentage of the labor force that is unemployed, the average earnings of employed workers, and so on, at a point in time in the census year.

The simplest approach to studying the relationship between immigration and domestic labor market conditions would be to collect cross-sectional

data and to look for relationships among the variables of interest, such as a negative correlation between the density of immigrants in a city and that city's wage rate. If cities with more immigrants are found to have lower wages, that finding would be consistent with the hypothesis that immigrants have a depressing effect on the local wage level.

Data from the 1990 census can be used to perform this exercise (U.S. Bureau of the Census 1993b). Taking the unit of observation to be the metropolitan area, and using as the sample the largest thirty metropolitan areas in the United States, one can compute the correlation between the average wage and salary income of households in the metropolitan area and the percentage of that metropolitan area's population that is foreign-born. Contrary to what the hypothesis of an adverse impact of immigration would predict, the correlation between these two variables is actually positive in that data.

Should this result lead us to the conclusion that immigration raises the income levels of the people in the cities to which the immigrants migrate? The answer is no. There are several problems with the most basic approach just outlined.

The first problem with the basic cross-sectional approach is that factors other than immigration that may affect wages are not taken into account. The conceptual basis of this first approach lies in the idea that variation in immigration is random, rather like a drug intervention in a medical study. By giving the drug to some patients and not others, and then comparing the health outcomes of patients who receive more or less of the treatment, the researcher can learn about the drug's effects. But if the people in the control group are systematically different from those in the treatment group (that is, there is nonrandom assignment), the study is invalid. Differences in the ex-post medical condition of the two groups of subjects may be due to their original heterogeneity rather than to any effect of the drug under evaluation.

In the study of immigration, the equivalent condition is that, along dimensions related to the outcomes of interest, cities that receive more immigrants must not be, *ex ante*, systematically different from those that receive fewer immigrants. Immigrants must be, in essence, "randomly assigned" to cities. There are many factors apart from immigration that determine the level of wages in a city. If immigration is correlated with these factors, the correlation between immigration

and wages may just be picking up the correlation between the third factor and wages. If these factors are observable variables on which the researcher has information, this bias can be eliminated by correcting for those factors (by including measures of them in a multivariate regression analysis). If the factors are unobservable, it is more difficult to obtain unbiased estimates. This problem is termed "omitted variable bias."

The second problem with the single cross-sectional approach is that of "endogeneity" (or "simultaneity"): in the relationship between two variables, the causality can run in both directions. If immigrants systematically choose to settle in locations with better labor market conditions, but their arrival causes a deterioration in those conditions, the sign of the resulting correlation between those two variables will be ambiguous. It will represent a combination of both phenomena and will not be amenable to a causal interpretation.

The Multiple Cross-Sectional Approach An improvement on an approach using cross-sectional data from one year is an estimation strategy that involves using two (or more) cross-sections of data. By using two point-in-time observations on the same set of geographic units, it is possible to circumvent some of the problems just pointed out.

Rather than correlating the stock of immigrants with the level of wages in a city, the fundamental idea behind this second approach is to look at the relation between the change in the stock of immigrants (that is, the net inflow) between two points in time and the change in the wage level between two points in time. In other words, it examines the correlation between changes in variables rather than levels of variables. This approach will go some way toward avoiding both the omitted variables bias and the endogeneity problem.

This method will help with the omitted variables bias if the omitted variables do not change over time and are hence subtracted away when the problem is considered in terms of changes in variables rather than in levels of variables. Such omitted variables may be referred to as "fixed effects." An example of something we think affects wages that does not change over time is the weather. Firms in cities with good weather and natural amenities are thought to be able to pay lower wages to their workers than firms in remote or unpleasant locations. Paying workers more to induce them to live in a place with less pleasing amenities is called a compensating differential. Since the weather does not change over time, however, we

expect weather to have no effect on changes in wages.

Might the weather factor lead to a spurious cross-sectional finding of immigrants raising wages? If immigrants care less about the weather than natives, there are likely to be more immigrants than natives in regions with bad weather—and hence high wages. This would lead to an apparent finding that immigrants raise wages. However, this correlation would not be due to any true positive impact of the immigrants, but only to the correlation between immigration and a factor that does have a positive impact on wages. In this example, using the naive cross-sectional approach would lead to an underestimate of the negative impact of immigration on wages.

The existence of interindustry wage differentials that are fairly stable over time might also cause wage differences across cities that are hard to control for in the cross-sectional analysis. Economists have long been aware of the fact that workers in some industries are paid more than in others. These differences appear to be for reasons unrelated to worker characteristics that affect productivity; they are more related to the economic rents, or profit levels, present in the industry. The auto industry is an example of a relatively high-paying industry. The city of Detroit has a heavy concentration of jobs in the auto industry for mostly historic and institutional reasons unrelated to current economic conditions or to immigration. Detroit is an example of a city with idiosyncratically high wages (holding constant the characteristics of the workers in that city). If immigrants move to cities with idiosyncratically high wages, the correlation between wage levels and immigration could well turn out to be positive even if the attributes of those cities, such as the educational level of their workforces, are controlled for. This positive correlation would lead researchers to reach the mistaken conclusion that immigration does not have a negative impact on wages.

Looking at changes between two cross-sections may also help solve the causality issue (endogeneity). If immigrants care only about the level of wages in a city, and not about the growth in wages, a correlation between changes in wages and changes in immigrant density will reflect causality running from immigrants to wages and not the reverse. This assumption seems strong, however, so the causality issue may not be fully solved by this technique.

A study that takes the multiple cross-sectional approach is that by Claudia Goldin (1994).

Goldin studies the impact of immigration in the United States around the turn of the century. She uses several cross-sectional data sets from the period 1890 to 1923, including city-level wage surveys and the decennial censuses of population from that period. Goldin studies the effect of a change in the fraction of a city's population that was foreign-born (immigrant density) on the change in that city's wage level over the same period, analyzing different time periods and different occupation and industry groups. The results vary for the different groups and time periods, but the most common result is that a one-percentage-point increase in immigrant density reduces wages by between 1.0 and 1.6 percent.

A new study by George Borjas, Richard Freeman, and Lawrence Katz (1997) uses the same approach to analyze decennial U.S. census data from 1960 to 1990. They find that the measured impact of immigration on wages is highly sensitive to the time period studied, as well as to the level of geographic aggregation chosen. Not only the magnitude but even the direction of the effect varies. These inconsistent results lead them to conclude that studies exploiting geographic variation in immigration are unreliable. Reasons for this will be discussed further later in the chapter.

Another study that takes the approach of examining changes in variables, rather than just their levels, is that by Robert LaLonde and Robert Topel (1991). LaLonde and Topel further improve on the naive single cross-sectional approach by using individual-level data on men. In addition to immigrant density, they include as explanatory variables in the analysis other city-level factors that plausibly affect wages, including any fixed effects. It would not be possible to identify and correct for city fixed effects in a single cross-sectional data set at the city level.

LaLonde and Topel (1991) study both levels and changes in the key variables of interest, using data from the 1970 and 1980 U.S. censuses of population. Their approach is to study the impact of immigrants on the wages of groups for which new immigrants might, *ex ante*, be expected to be the closest substitutes in the labor market. The largest effect they find is that increasing the density of new immigrants in a local labor market by 10 percent reduces the wages of other new immigrants in that area by 0.3 percent. The effects for young African Americans and Latinos are of the same size but are not statistically significant. The results are the same regardless of whether the authors use as their earnings variable weekly or an-

nual wages. This implies that the factor making these two variables differ—weeks of employment—is not significantly affected by immigration.

These two studies conclude that immigration is not a cause for concern about wages, but a number of imperfections remain in their approach. The Goldin study is affected by what may be called the composition problem. City-level wages are a composite of the wages of immigrants and natives in that location. If immigrants earn less than natives, cities with higher proportions of immigrants will have lower-than-average wages, even if immigrants have no negative impact on the native wage. This problem is well known and usually arises when the data do not allow immigrants and natives to be distinguished, hence precluding an examination of the outcomes for natives alone.

The remaining concern about LaLonde and Topel's paper is one that applies generally to papers relying primarily on cross-sectional variation in immigrant density (including Goldin's) and is related closely to the earlier discussion of trade and factor price equalization. This will be discussed in the next section.

The Issue of Factor Price Equalization When using the geographic approach, even if immigration would have a strong impact on employment and wage levels in a city, such an effect may not be detectable when labor markets are linked. An analogy can be made to a pool of water. If a bucket of water is poured into the pool, the water level at that particular spot will not be higher than the water level in the rest of the pool. Using a geographic, or cross-sectional, approach would lead to the conclusion that pouring water into a pool does not affect the amount of water it contains. This approach would miss the fact that the overall water level of the pool had risen.

Applying this analogy to the labor market, because goods and factors of production can (at least in theory) flow freely across the United States, we might expect that high wages in one geographic area would cause workers to flow in or firms to flow out, bringing the wage level back into line with wages elsewhere. The same would apply to the prices of other factors of production or to the unemployment rate.

As already explained in the previous section, economic theory predicts that, in the face of flows of goods, labor, and capital, prices will tend toward equality across different geographic areas. Although there is no strong evidence of factor price equalization across countries, it is more probable

that it holds within countries. If the tendency toward factor price equalization is strong, the geographic approach will miss much of immigration's impact by focusing on differences across areas rather than on the aggregate.

One way of studying whether flows leading to factor price equalization nullify the cross-sectional approach is to look at one example of such a flow, namely, the migration of labor within the country. If immigration to a particular location leads to offsetting migration flows within the country, then, as in the analogy to the pool, the geographic approach will not pick up immigration's full impact.

Studies of internal migration in the United States are mixed in their conclusions about the reaction of internal migration to immigration from abroad. Two papers, by Randall Filer (1992) and by Michael White and Lori Hunter (1993), use data from the 1970s and conclude that internal migration does indeed offset inflows from abroad, tending to equalize population growth across areas. William Frey (1995d) confirms this finding for migration by low-income people, using 1990 census data, as do George Borjas, Richard Freeman, and Lawrence Katz (1997) using data from the 1960 to 1990 censuses. Timothy Hatton and Jeffrey Williamson (1995) also find evidence of this in data from the turn-of-the-century United States. They find that in the period 1880 to 1910 immigrants predominantly arrived in the Northeast, pushing natives to the West. However, other work—by Kristin Butcher and David Card (1991), Michael White and Zai Liang (1994), and Card (1997)—using recent U.S. data, does not find evidence of offsetting migration flows. Two studies of the European experience—Jörn-Steffen Pischke and Johannes Velling (1997) looking at Germany and Jennifer Hunt (1992) studying France—also do not find that internal migrants avoid areas of high immigration.

Other, more direct studies of factor price equalization indicate that, contrary to the example of the pool of water, the process of equalization across the aggregate labor market is quite slow. A recent study by Jörg Decressin and Antonio Fatàs (1995) finds that shocks to labor market conditions across regions in Europe take about four years to dissipate. In the United States, Olivier Blanchard and Lawrence Katz (1992) find that labor market shocks persist for even longer, with the local effects on unemployment and labor-force participation remaining for six years and the effects on wages persisting for as long as a decade. These durations indicate that the problems posed to the

geographic approach by factor price equalization are not very serious.

The Instrumental Variables Approach The major problem remaining for cross-sectional studies using data at a level other than the individual level (that is, most studies, which use geographic units of observation) is that of causality. Immigrants' choice of location may be based not only on the level of the wage but also on the growth in the wage, so that differencing will not eliminate the causality problem. In some cases, a statistical solution called instrumental variables estimation may be used to establish causality.

There are many questions in economics in which the empirical identification of a particular causal effect is difficult because the observed relationship between two variables of interest involves more than one direction of causality. If variable X affects the level of variable Y , but variable Y also affects the level of variable X , it is difficult to identify separately the two effects that lead to their correlation. One way to isolate the effect of X on Y is to find a variable that is correlated with variable X but uncorrelated with any factor affecting variable Y . Call this variable Z . Using the instrumental variables estimation technique, the correlation of Z with X can be used to estimate the effect of X on Y . A variable such as Z is called an instrumental variable or "instrument."

In the case of immigration, the identification problem arises from the possibility that two connected phenomena are occurring. First, the more immigration there is to a city, the more its wage level falls. Second, the more the wage in a city falls, the less immigration there is to that city. In order to test the first proposition, it is necessary to find a variable that is correlated with immigration to a city (that is, the change in the immigrant density) but does not itself influence that city's wage level and is not correlated with other omitted factors that do (which would cause a simultaneous feedback to changes in immigration).

Using instrumental variables estimation to correct the endogeneity problem that may exist in other studies has a potential second beneficial effect on the estimation as well. If the instrumental variable, in addition to being uncorrelated with nonimmigration factors affecting wages, is also uncorrelated with the flows of capital, labor, and goods that lead to factor price equalization, the instrumental variables approach will get around the problem of factor price equalization posed by the other approaches as well.

Because so few variables in economics can really be considered exogenous (that is, not determined within the system being studied), it would very often be desirable to use the instrumental variables technique to solve the identification problem in estimation. However, much empirical analysis calls for instrumenting for the same reason that it is difficult to identify an instrumental variable in any given estimation situation: it is hard to find a variable that is correlated with X and yet uncorrelated with Z , except through its correlation with X .

One study that identifies such an instrumental variable is that by Joseph Altonji and David Card (1991). Altonji and Card set out to study the effect of changes in the immigrant density of a city on the change in that city's labor market outcome variables. The instrument needed in this case is one that is correlated with the change in the city's immigrant population but does not itself influence the change in labor market outcomes. The instrument they identify is the density of immigrants in that city in the base year. Ann Bartel (1989) found that in the 1970s immigrants tended to settle in cities where other similar immigrants already lived. Thus, the original immigrant density of a city is a good predictor of future immigration to that city. Assuming that the original ethnic makeup of a city in a base year does not have an effect on changes in the labor market from that base year to a later year, the fraction of a city's population who were foreign-born in the base year can be used to instrument for the change in that density over time. If these assumptions are correct, this study avoids all the problems mentioned.

Altonji and Card (1991) use data from the 1970 and 1980 U.S. censuses. They focus on the groups thought most likely to face competition in the labor market from immigrants: white male high school dropouts, and African American and white females, and African American males with a high school education or less. They study several different labor market outcomes. Their results suggest that immigrants had an unexpected negative effect on unemployment in the census week, the opposite of the expected result. Immigrants also had a negative effect on the fraction of the population who worked in the previous years and on weekly earnings in the previous year. No significant impact was found on the labor-force participation rate, the employment-to-population ratio in the census week, or the number of weeks worked in the previous year. The magnitudes of the coefficients imply that a one-percentage-point increase in the percentage of foreign-born in a city

reduces the unemployment rate by 0.23 percentage points; reduces the number who worked in the previous year by 0.25 percentage points; and reduces wages of unskilled natives by 1.2 percent, at most. Put in terms of an elasticity, a 10 percent increase in the foreign share of the population implies a 0.86 percent fall in weekly earnings (wages).

The technique developed by Altonji and Card (1991) is applied to data from West Germany by Pischke and Velling (1997). Their data cover changes over the period from 1985 to 1989. This study finds no significant adverse effect of immigration on employment or unemployment.

A new study by Card (1997) divides workers into ten skill groups and, using the 1990 census, estimates the effect of changes in the relative supplies of different skill groups caused by immigration across cities on employment and wages, finding small effects.

It is sometimes difficult to decide without some point of comparison whether the size of an effect is large or small. For the Altonji and Card (1991) result, an interesting comparison is with the effect of "generational crowding," studied by Finis Welch (1979). Welch sought to estimate how much the earnings of baby boomers were lowered by the large numbers in their age group or cohort. The theoretical basis is that people of different ages are imperfect substitutes in production, and that we can apply the same analysis to the case of an increase in fertility that we did to an increase in the number of unskilled workers, for example. We expect that a shift out in the supply of an age group will lower the wages of that age group. Welch finds that an expansion in the cohort size of 10 percent had an effect on weekly wages of white male high school dropouts very similar to the immigrant effect calculated by Altonji and Card. His finding that the effect on annual wages is considerably larger suggests that an important part of generational crowding comes through less employment, while Altonji and Card found small or positive effects on employment outcomes. Furthermore, the actual increase in cohort size induced by the baby boom is calculated by Welch to have reduced high school dropout wages by 12 percent between 1967 and 1975. To cause the same wage reduction would require a 140 percent (or ten-percentage-point) increase in the immigrant share, using Altonji and Card's estimate.

Using Production Functions The production function is a mathematical formula that expresses

the quantity of output as a function of the quantity of the various inputs to the production process (namely, factors of production). Data on inputs and output can be used to estimate empirically the form and parameters of production functions. These parameter estimates reveal the relative contributions of different factors of production as well as on what the marginal contribution of each factor depends. They show the rate at which it is possible to trade off one input for another in production, what is known as the elasticity of substitution among factors. If we make the common assumption that factors are paid according to their marginal contribution to output, the estimated production function parameters can be used to calculate the effect that a change in the quantity of one factor would have on the market price of another factor.

Applied to the case of immigration, immigrants and natives are allowed to represent two distinct inputs to the production process. Empirically estimating a production function with native and foreign-born labor as two of the inputs allows estimation of the effect on the wages of the native-born of an increase in the number of foreign-born. In theory, this wage impact could be positive or negative, depending on the relationship between these two factors of production.

A 1982 paper by Jean Grossman uses this approach to gauging the wage impact of immigration. Grossman estimates the form of the production function using cross-sectional data from 1970 on the inputs and outputs of SMSAs across the United States. Her estimates suggest that a 10 percent rise in the number of immigrants employed in a city lowers the wages of natives in that city between 0.2 and 0.3 percent. In a similar study of Canada for 1980, Ather Akbari and Don DeVoretz (1992) find virtually no effects and conclude that immigrants do not have an impact, except possibly in those industries in which they are most heavily concentrated.

George Borjas (1987b) performs the conceptually closely related exercise of estimating labor demand functions using SMSA-level data from the 1980 census. He finds that a 10 percent rise in the number of immigrants reduces the wages of natives by at most 0.4 percent. (Both Borjas and Grossman find that the wages of earlier immigrants are more strongly affected.) This is about half the size of the Altonji and Card (1991) estimate.

Using the 1980 and 1990 censuses, David Jaeger (1996) estimates the elasticity of substitution

between immigrants and native workers for various educational groups and concludes that within educational groups immigrants and natives are nearly perfect substitutes. He then estimates an aggregate production function and finds substantial effects: immigration explains up to one-third of the decline in the wages of high school dropouts during the 1980s.

Exploiting the Time Dimension

Some researchers have used greater—or in some cases total—reliance on time series variation in data to circumvent the thorny issues associated with factor price equalization and endogenous immigrant location choice present in more geography-based analyses. Pure time series data consist of observations at different points in time of a single entity, usually a country. “Panel” data combine cross-sectional and time series variation by following several entities—which may be individuals, firms, or countries—over time. If time series data on the relevant variables, such as wages and immigration, are used at the national level, the estimate of the impact of immigration comes from seeing how changes in immigrant flows over time are correlated with changes in wages over time. The estimate therefore is not influenced by where in the country immigrants settle, nor by the fact that the impact of immigrants may be spread around the country. However, time series estimation has its own endogeneity problem: more immigrants will choose to go to (or will be allowed to enter) a country at times when labor market outcomes are favorable in that country. Thus, time series will also tend to underestimate the impact of immigrants unless the instrumental variables technique is used.

Time Series David Pope and Glenn Withers (1993) use pure time series analysis for Australia over the period from 1861 to 1981 and find no negative effects of immigration on native unemployment or wages. In their study, they use the labor market characteristics of sending countries and transport costs to Australia as instrumental variables for the migration rates. These variables are likely to influence migration rates but are unlikely to influence Australian labor market outcomes in other ways.

Hatton and Williamson (1995) use time series data for the United States between 1890 and 1913 and find that a one-percentage-point increase in the proportion of immigrants in the labor

force reduced wages by 0.4 to 0.5 percent. This is close to half the magnitude found by Goldin (1994) (and by Altonji and Card [1991]), although, given the huge immigrant flows of this period, it still translates into substantially lower wages than in the absence of immigration. Because of constraints on the historical data, the results are affected by the composition problem.

Wage Inequality Literature A series of recent papers motivated by rising inequality in the United States has put more weight on time series variation in immigration rates. A very large number of papers have sought to document and explain the rise in U.S. wage inequality, a level of inequity which is high by international standards (for a summary, see Levy and Murnane 1992). The ratio of the wage at the ninetieth percentile of the distribution to the wage at the tenth percentile of the distribution (a simple measure of inequality), for example, rose by almost 30 percent from the late 1960s to 1987 (Katz and Murphy 1992) and was about 28 percent higher than in Britain in 1989 (Katz, Loveman, and Blanchflower 1995). Much attention has been given to the fact that in the 1980s a contribution to the rise in overall inequality was made by an increase in the return to education—that is, by an increase in the wage gaps between people of different educational levels. The absolute wages of high school dropouts fell over this period, while those of college graduates rose considerably. One hypothesis tested is that immigration of largely unskilled workers may have played a role by increasing the supply of high school dropouts. The usual principle of supply and demand applies here: if the number of high school dropouts rises relative to the number of college graduates, the relative wage of high school dropouts would be expected to fall.

The papers testing this hypothesis have estimated elasticities similar to or smaller than the Altonji and Card (1991) and Goldin (1994) estimates, but the context of wage inequality in which the results are set suggests that immigration may have economically important effects. Such studies have used both national-level time series and a panel of cities and use yearly data from the Current Population Survey (CPS).

Because the variable of interest in these studies is a relative rather than an absolute wage, an assumption is necessary to compare these results to the others discussed. We assume that falls in the relative wage of high school dropouts compared to that of college graduates come entirely from a fall

in the high school dropout wage. In fact, some comes from a rise in the college graduate wage, and hence these calculations overstate the negative impact of immigration on the absolute wage of the unskilled.

Borjas, Freeman, and Katz (1992, 1997) use time series data for the United States from 1967 to 1987 using the following estimation strategy. The main variable investigated is the ratio of average college graduate wages in a given year in either the United States or a particular region to average high school dropout wages. A regression is run to establish the influence on this ratio of the relative supply of different educational groups—that is, how the number of high school dropouts or high school graduates relative to college graduates influences the ratio. The studies find as expected that increasing the relative supply of high school dropouts reduces their relative wage. A calculation is then made as to the contribution of immigrants to the high school dropout group in the 1980s, and hence as to the contribution of immigrants to the fall in the relative wage of high school dropouts.

Between 1980 and 1988 immigrants as a proportion of the labor force rose from 6.9 percent to 9.3 percent, and the results of Borjas, Freeman, and Katz's first paper (1992) suggest that over the same period immigration accounted for one-quarter of the 10 percent decline in the relative earnings of high school dropouts. Using the figures on change in immigrant density, and making the assumption that the relative wage change came entirely from a fall in the dropout wage, we can convert this number to make it comparable to those in the studies discussed earlier. Thus, a one-percentage-point increase in the proportion of immigrants reduces the absolute wage of dropouts by at most 1.2 percent, the same magnitude as calculated by Altonji and Card (1991). Borjas, Freeman, and Katz's second study (1997) finds that on the order of half of the decline in the wages of high school dropouts from 1980 to 1995 can be attributed to immigration.

George Borjas and Valerie Ramey (1995) look at relative wages of high school dropouts and college graduates in the United States, using a panel of forty-four SMSAs from 1977 to 1991. Their estimates suggest that a one-percentage-point increase in the fraction foreign-born reduces the wage of high school dropouts relative to college graduates by 0.6 percent. This is half as large as the Altonji and Card (1991) estimate.

Although time series analysis avoids the cross-

sectional biases due to immigrant location choice and factor price equalization, data constraints cause the papers emphasizing the time series dimension to introduce new biases, leading to an overstatement of the impact of immigration. The problems arise from the fact that immigrants (and hence immigrant density) can be identified in the CPS only in certain years. This means, first, that in the central regression the wages being measured are those of both immigrants and natives, and hence the regression suffers from the composition problem: it does not measure the effect of the size of educational groups on the wages of native high school dropouts relative to native college graduates, but rather of all high school dropouts to all college graduates. The same problem exists with the independent variable used in the regression (in the case of Borjas, Freeman, and Katz [1992, 1997]): using the relative supplies of all high school dropouts and all college graduates constrains the impact of a native high school dropout to be the same as that of an immigrant high school dropout. A similar comment applies to the findings of Hatton and Williamson (1995), who in their regression estimate the impact of the total labor force on wages.

For the papers using recent data, there is good reason to expect the composition effect to be large and the impact of immigrant dropouts on natives to be lower than the impact of native dropouts on natives: immigrant and native dropouts are unlikely to be good substitutes for each other because even among high school dropouts immigrants have much less education. For males in 1989, 30 percent of U.S.-born high school dropouts had eight or fewer years of education, while 75 percent of foreign-born high school dropouts had eight or fewer years (LaLonde and Topel 1994). A calculation can be made as to the bias introduced by the composition effect: if foreign-born high school dropouts earn 20 percent less than native dropouts, the increase in their number from 1980 to 1988 would be expected to reduce the average dropout wage by 1.5 percent because of the composition effect alone, and this magnitude is large compared to the total impact of 2.5 percent ascribed to immigration in Borjas, Freeman, and Katz (1992, 1997).

Borjas and Ramey (1995) attempt to deal with the second problem by including the percentage foreign-born explicitly on the right-hand side of their regression; doing so avoids constraining the impact of natives to be the same as that of immigrants. However, their foreign-born variable is

constructed by interpolating between the 1970 and 1980 censuses and the 1989 CPS, and the time series variation is thus to a large extent a trend.

Growth Studies An empirical assessment of the effect of immigration on growth raises a new set of measurement problems. The principal one is simply that while immigration may affect growth, growth surely affects immigration: at least when choosing among countries with similar levels of economic or labor market well-being, an immigrant will prefer a country with strong growth, which promises higher future levels of well-being than a country with slow (or negative) growth. As always with statistical simultaneity problems, if a good instrumental variable can be found, causality may be established. In this case, a variable that influences immigration but does not directly influence growth is required.

An empirical examination of the impact of immigration on growth is provided by Robert Barro and Xavier Sala-i-Martin (1992) for Japan and the United States. The main thrust of this paper is an issue of current interest to macroeconomists seeking to explain relative standards of living in different countries: whether the per capita income of regions or countries tends to converge over time. They use panel data on regions over time to run regressions, including the level of per capita income and net migration rates as predictors of the growth in per capita income. The coefficient on the level of per capita income is a measure of convergence, or regression to the mean: convergence occurs if regions with currently low levels of income grow faster and countries with currently high levels of income grow slower. The authors assess the effect of immigration not only by the coefficient on immigration but also by its effect on the convergence coefficient. For both Japan and the United States, adding migration to the convergence regressions raises the convergence coefficient slightly and yields a positive coefficient on migration. For the United States, this coefficient suggests that a one-percentage-point higher net migration rate is associated with a growth rate increase of 0.1 percent. These results are for a regression without use of instrumental variables.

Barro and Sala-i-Martin (1992) then propose the use of instrumental variables for migration: temperature and population density of the region. Temperature would be expected to influence migration, but not economic growth, directly, while the idea behind population density is that it is a

proxy for housing costs, or possibly for congestion, which would discourage migration but may not directly affect economic growth. When the instrumental variables specification is used, the coefficient for migration becomes statistically insignificant for both the United States and Japan (and the convergence coefficient is virtually unchanged). The point estimate suggests that a one-percentage-point rise in the net migration rate raises long-term growth by 0.01 percent for the United States, and by 0.04 percent for Japan, but statistically it cannot be asserted with confidence that these numbers are different from zero. The authors conclude that migration has little effect on growth. However, the results look as though the instruments may not have been suitable: either they may not have been very highly correlated with migration or they influenced growth through a channel other than migration.

A puzzle arises when the work of Barro and Sala-i-Martin (1992) is compared with that of Blanchard and Katz (1992). The latter study uses data on U.S. states to examine how states adjust to economic shocks. If a state experiences a shock to labor demand not necessarily experienced by the whole country—for example, Texas is hit by a fall in the price of oil—one adjustment mechanism would be for the wage to fall, so as to keep everyone who wished to work at the lower wage employed. This would imply a fall in the labor-force participation rate. Another possibility is that the wage might not fall quickly: in this scenario, employment falls by more than in the first case, labor-force participation falls less, owing to the high wage, and unemployment arises. Finally, rather than remaining unemployed, some people may prefer in the short or long term to leave the state.

Blanchard and Katz (1992) estimate a system of equations for employment growth, participation rates, and unemployment rates in states over time. They find that after a shock to a state's employment growth, participation and unemployment rates eventually return to their preshock levels. Employment returns to its old growth rate, but no higher; as a result, employment is permanently lower than it would have been in the absence of a shock. (Since the growth rate was lower for a while, it would have to rise above its old rate to bring the level of employment back to where it would have been.) Since participation and unemployment rates are back at their old levels, the lower employment must be due to emigration from the state. This picture suggesting that migration is a very important equilibrating mechanism

for recovery to shocks to a state is at odds with the Barro–Sala-i-Martin results.

Currently we have only international evidence to turn to in order to shed light on this inconsistency. Juan Dolado, Alessandra Goria, and Andrea Ichino (1993) use a panel of countries to study migration among the Organization for Economic Cooperation and Development (OECD) countries. They base their study on the theoretical growth model described in the theory section by performing what is known as structural estimation, which involves entering the variables into the regressions in nonlinear ways determined by the equations of the theoretical model. Structural estimation, while following theory more closely than the usual linear reduced-form alternative, does not obviate the need to use instrumental variables in cases where simultaneity is an issue. These authors use as their instruments for migration lagged values of migration (the value of migration in the previous period), savings rates, schooling rates, and population density. This paper is also primarily concerned with the convergence issue and therefore does not report the coefficient on the net migration variable. However, the finding that adding migration reduces the convergence coefficient is more in accordance with the Blanchard and Katz finding.

As discussed earlier, the results we should expect here are determined by the human-capital level of the migrants compared to the natives. Dolado, Goria, and Ichino (1993) document that immigrants to OECD countries have lower human capital than natives. George Borjas, Stephen Bronars, and Stephen Trejo (1992) investigate empirically what influences the migration of high- and low-skill individuals within the United States. A theoretical model called the Roy model makes predictions about which sorts of people will migrate from one country to another. Clearly people move from lower to higher wage or income regions, but this calculation is affected by the position of the individual in the distribution of the original region and by their likely position in the distribution of the potential destination. A person with high skills who lives in a region with a compressed wage distribution (where wage gaps between high- and low-skilled individuals are low) will have more to gain from leaving than an unskilled person in the same region and will seek out a region with a dispersed wage distribution (as well as one with a high mean wage). An unskilled person may still want to leave a region of compressed wages if the mean wage is low, but such a person, in addition

to seeking a region with a high mean wage, will tend to choose one with a compressed wage distribution.

Borjas, Bronars, and Trejo (1992) find empirical evidence that, conditional on the mean wage, regions with compressed wages encourage the skilled to emigrate, while regions with dispersed wages attract the skilled. This theory and empirical evidence show that the quality of immigrants compared to natives will depend on the dispersion of wages in the receiving country. Migrant quality has not been explicitly accounted for thus far in the internal migration and growth literature. It is possible that doing so might help resolve the puzzle of U.S. internal migration and its effects on growth.

Exploiting Natural Experiments

Certain studies have focused on episodes of immigration that the authors hope suffer less from the factors that make statistical analysis, especially the determination of causality, so problematic. These studies focus on an episode of immigration that was politically motivated and hence does not suffer from the time series problem that immigrants arrive at times when the economy and labor market are doing well. In some cases, the choice of location within the receiving country was also determined to a large extent by noneconomic factors, such as proximity to the point of arrival, which in turn was related to the political event that caused the arrival. Another advantage is that the episodes chosen were very sudden and large inflows that occurred within a brief space of time; focusing on such events should make it easier to isolate responses to the arrivals from other factors that change more gradually over time.

One example of such an episode is the arrival in Miami of the Cuban refugees of the Mariel boatlift in May 1980, examined by David Card (1990). Most of the refugees went to Miami because it is close to Cuba; they increased the population of Miami by about 7 percent. Since both timing and location were fairly exogenous, Card simply examines the evolution of variables of interest in the period immediately following the boatlift, comparing the outcomes with those in similar cities and ascribing differences in patterns to the arrival of the Mariels. Card uses yearly data from the Current Population Survey to follow wages, employment rates, and unemployment rates for unskilled whites and African Americans, for non-Cuban Latinos, and for Cubans. Only Cubans appear to

have been negatively affected. It is not possible in the data to distinguish Mariel Cubans from earlier Cuban arrivals, however, and thus the observed decline of Cubans could reflect assimilation difficulties (or lower quality) of the Mariel Cubans. Some information on this point is available from a supplement to the Current Population Survey: it shows that the Mariels did indeed earn less and had higher unemployment than other Cubans, by an amount that is consistent with the "composition effect" being the cause of the observed overall decline for Cubans. Earlier Cubans thus seemed to have been unaffected by the Mariel boatlift. Card finds that population growth in Miami slowed, however, and he speculates that other migration to Miami may have been reduced in response to the boatlift. Since Card examines yearly data, any type of factor price equalization must have happened very quickly.

A second example of such a study is that by Jennifer Hunt (1992), which examines the repatriation to France in 1962 of Algerians of European origin. The independence of Algeria from France, agreed upon in March of that year, prompted the emigration within the space of a year of virtually the entire population of European origin, most of whom went to France. This inflow increased the French labor force by 1.6 percent, but the repatriates were very unevenly distributed through the country: a vast majority settled in the south, where the climate was more similar to Algeria's and a small number of people who had left Algeria during the war of independence had settled. Hunt uses this cross-sectional variation in the 1962 and 1968 census data to assess the impact of the repatriates on the labor market outcomes of natives and addresses the usual cross-section biases by instrumenting the fraction of immigrants with the temperature and the density of pre-1962 repatriates. She finds that a one-percentage-point increase in the repatriate share of the labor force reduced the wage of a region by at most 0.8 percent (zero in some specifications) and raised the unemployment rate of natives by 0.2 percentage points. Hunt found that internal migration within France did not respond directly to the location choice of the repatriates (but did respond insofar as the repatriates affected labor market variables), and that in fact foreign (nonrepatriate) immigrants appeared to be attracted to areas with more repatriates.

In a recent paper, Rachel Friedberg (1997) studies the mass migration of Russians to Israel in the early 1990s. This migration was precipitated

by the lifting of emigration restrictions in an unstable Soviet Union and by the open immigration policy of Israel toward Soviet Jews, who faced more restrictive entry policies elsewhere. The immigration was therefore largely exogenous to economic conditions in Israel. Using both individual and occupation-level data, the paper exploits the fact that the occupational distribution of the Russian immigrants was different from that of the native population. There is a negative cross-sectional relationship between native wages and the presence of Russian immigrants in an occupation. There is a weaker negative relationship between native wage growth and the influx of immigrants. This demonstrates that some of the cross-sectional relationship is due to the fact that immigrants entered low-wage occupations, rather than being fully attributable to a depressing effect of immigration on wages. In addition, since the distribution of immigrants across occupations may not have been exogenous to relative wage or employment growth across occupations, the paper instruments for the influx of immigrants into an occupation, using information on the immigrants' previous occupations in the Soviet Union. When the instrumental variables technique is used, the negative impact disappears. Immigrants entered low-wage, low-wage-growth, contracting occupations but did not have an adverse impact on the labor market outcomes of natives.

A final example of the natural experiment literature is a paper by William Carrington and Pedro deLima (1996) that examines the return of Portuguese colonialists from Africa after a Marxist revolution in 1974 led to the sudden independence of the colonies. They use several approaches to the problem: time series analysis, comparisons with Spain, and cross-sectional analysis across the twelve provinces of Portugal. The approaches give quite different answers, making it difficult in this case to reach a firm conclusion. Apart from this paper, the natural experiment literature adds to the evidence suggesting a limited impact of immigrants on natives.

Computable General Equilibrium

A series of studies of immigration use the computable general equilibrium technique. This technique uses a set of equations describing the whole economy and sets their parameters to numerical values obtained from estimation. Data on immigration can be fed into an appropriate computable

general equilibrium model, and predictions of the responses of other variables in the economy can be obtained. These studies (see, for example, Hatton and Williamson 1995; O'Rourke, Williamson, and Hatton 1994; Williamson 1974c; Williamson and Lindert 1980) suggest that the large immigrant flows of the nineteenth century had an important negative impact on real wages, leading to convergence with the old world and higher inequality.

CONCLUSION

Economic theory predicts variously that immigration may be beneficial or detrimental to the receiving country's workers, or indeed that immigration may have no effect on them. Even the sign of this impact must therefore be determined empirically. Given the widespread nature of the popular view that immigration has large adverse effects on the economic outcomes of the native-born population of the United States, there is surprisingly little evidence to support this. There are several difficulties associated with empirical estimation of the impact of immigration, and we have argued that some studies may over- or underestimate it. Nevertheless, it appears to be difficult to detect large immigration-induced deteriorations in labor market outcomes, while the evidence on the link between immigration and economic growth is for the moment less certain. Most research finds that a 10 percent increase in the fraction of immigrants in the population reduces the wages of even the least skilled native-born workers by at most 1 percent. This magnitude would be large enough for immigration to have been responsible for one-quarter of the rise in wage inequality in the United States in the 1980s. We have argued, however, that the true magnitude is likely to be considerably smaller. Evidence of immigrants reducing employment or labor-force participation rates or increasing the unemployment rate is even harder to find.

In evaluating the overall consequences of immigration for the United States, other factors not considered here—both economic and noneconomic—must be taken into account. These include the effect of immigration on government revenue and expenditures, consumer prices, trade, residential patterns, crime, and the demographic composition of the population, to name a few. The social dimensions of immigration may in fact play a more important role than economic factors in shaping public opinion on immigration. A re-

cent, well-publicized report by the National Academy of Sciences (Smith and Edmonston 1997) concludes that the overall economic gains and losses from immigration are both modest: the domestic gain is between \$1 billion and \$10 billion

per year, in the context of an \$8 trillion U.S. economy. Future interdisciplinary research could usefully link economic and social considerations and examine the interaction of both with public opinion.