Editors' Summary

THE BROOKINGS PANEL on Economic Activity held its seventy-second conference in Washington, D.C., on September 6 and 7, 2001. This issue of *Brookings Papers on Economic Activity* includes the papers and discussions presented at the conference. The first paper revisits the economy of the former East Germany ten years after reunification, seeking to explain why its convergence toward western German economic performance has stalled. The second paper reviews macroeconomic policymaking in Japan to determine whether the stagnant Japanese economy has responded to policy stimulus as conventional models would predict. The third paper tests alternative explanations of the greater generosity of social welfare systems in Europe than in the United States. And the concluding paper attempts to reconcile the observed high return on equities with reasonable levels of risk aversion on the part of individuals.

THE REUNIFICATION OF GERMANY, which began with a bloodless revolution in 1989, was one of the most dramatic economic and political events of the twentieth century. The process of integration that began officially in 1990 with economic, social, and monetary union provides a vivid case study in the difficulties, even under highly favorable conditions, in achieving the convergence of living standards in two regions at very different levels of development. The less developed former East Germany faced no language barrier; it could easily import the political, legal, and economic system of the west; its people could migrate freely; and it received massive inflows of capital. In the first paper of this issue, Michael Burda and Jennifer Hunt examine in detail the economic performance of eastern and western Germany during the first decade after reunification, analyzing changes in output, employment, investment, consumption, and wages to better understand the productivity gains of the east.

The authors' overview of economic developments since reunification identifies a number of surprises, both good and bad. After an initial drastic decline, eastern Germany's GDP per capita grew rapidly from 1992 to 1995, reaching two-thirds of western levels. Labor productivity, consumption per capita, and wages did even better, rising to more than 70 percent of western levels by mid-decade. The authors illustrate the convergence in consumption by comparing rates of household ownership of selected consumer durables. In 1991 few eastern Germans owned dishwashers, microwave ovens, video cameras, refrigerator-freezers, or even telephones. But by 1998, 44 percent of eastern German households had dishwashers, 54 percent had microwaves, 38 percent had video cameras, 35 percent had refrigerator-freezers, and 97 percent had telephones. By that year the prevalence of ownership of automobiles, telephones, color television sets, refrigerators, refrigerator-freezers, stereos, video cameras, and washing machines was essentially the same for eastern and western Germans.

This convergence is impressive and happened much faster than previous studies of regional convergence would have predicted. Viewed from the end of the decade, however, prospects for complete convergence do not seem as bright. Since 1995, consumption and GDP per capita, wages, and labor productivity have all been stuck at essentially the same fraction of their western counterparts. And whereas output and consumption have more than regained the ground lost following the initial shock of the transition, the labor market has yet to recover. Various measures give different indications of the extent of unemployment: survey measures suggest that unemployment averaged 13 percent from 1994 to 1999; official estimates are higher, with unemployment approaching 20 percent in 1998, or 27 percent if estimates of hidden unemployment are included. The share of the eastern working-age population that is employed declined from 83 percent in 1990 to 65 percent in 1999.

To help explain the early successes and the more disappointing performance that followed, the authors undertake a careful analysis of eastern Germany's output growth, grounding it in a conventional growth accounting framework. They believe much of the large gap in labor productivity between east and west in 1990 reflected the relatively low capital-labor ratio in the east. Given political stability and firmly established rules for a pan-German market economy, this difference in the capital-labor ratio provided a powerful incentive both for labor migration to the west and for capital accumulation in the east, the latter from both capital inflows and

internal accumulation. The authors use a simple model to show that the configuration of the German economy in long-run equilibrium, when factor returns have been equalized, will depend on the importance of each of these channels during convergence. If labor migration to the west is slow and capital inflows to the east are large, the eventual distribution of output and employment will be roughly proportional to the initial populations in east and west. If instead labor is highly mobile but capital is slow to move, output and population in the east will be much smaller in equilibrium. And if capital inflows are small enough, internal saving in the east could be so low that net investment would be negative, and the east could largely empty out.

The authors go on to analyze what has actually happened thus far. In assessing the importance of capital formation in the east during the last decade, they focus on investment in equipment. They believe that this, rather than structures, is the key bottleneck to development, and that data on structures are, in any case, a poor measure of capital services provided. Cumulative net investment in equipment during 1991–98 was impressive, totaling DM 504 billion at 1995 prices, roughly equivalent to current real GDP in the eastern states. By comparison, cumulative net equipment investment in the west during the same period was about two-thirds of its GDP. In contrast to this rapid growth of the capital stock, the labor force in the east actually declined after reunification. From 1988 to 1999 the registered population in eastern Germany fell by nearly 9 percent; net emigration to western Germany and an excess of deaths over births were the two main sources of the gross decline (accounting for 7.2 and 4.7 percentage points, respectively; these were partly offset by an influx of migrants from abroad). The time pattern of migration itself is fascinating. East Germany suffered outflows in 1989 and 1990 of 2.5 percent of its population. Emigration soon fell to about half that rate, but immigration to the east, including returnees, also rose, leaving only modest net population outflows after the mid-1990s. As would be expected, the young have been the most likely to move. Among the population that remained, labor force participation rates declined between 1991 and 2000, by 6 percentage points for men and by 5 percentage points for women.

Burda and Hunt calculate that, even assuming an initial equipment stock of zero, accumulation had led by 1998 to an equipment capital–output ratio in the east of roughly 75 to 90 percent of that in the west. As Janet Yellen shows in her discussion of the paper, a capital shortfall of the

magnitude suggested by these numbers would result in only a modest shortfall in labor productivity if the aggregate production functions in east and west were identical. Hence much of the remaining difference between eastern and western labor productivity is likely to reflect lower total factor productivity (TFP).

Applying the standard Solow decomposition of output growth since 1991, the authors find that TFP growth in the east was strikingly higher in the first than in the second half of the decade. Output grew at a remarkable 6.8 percent annual rate during 1992–95 despite a decline in employment that contributed a negative 1.8 percent a year to output growth. The rapid growth of the capital stock more than offset this decline, with capital contributing 4.2 percent a year. This leaves TFP rising by an extraordinary 4.4 percent a year. The authors estimate that, following this surge, TFP actually declined by about 1 percent a year during 1995–99. This development appears systematic, affecting all the eastern states in the same qualitative fashion. At the same time, TFP growth in the western states, which had been essentially zero during the early years of integration, returned to slightly over 1 percent during the later period, close to its historical norm.

The authors note that a large portion of the initial surge in productivity and wages may have nothing to do with these customary sources of growth but instead represented movement toward the efficient production frontier. Some firms reduced employment by as much as 80 to 90 percent, suggesting that some of the productivity surge came from firms discharging redundant labor and from other profit-driven reorganization. If these potential sources of productivity growth were fully exploited in the early part of the decade, they help explain the lack of convergence in TFP since 1995 and imply that further convergence will be hard to achieve.

The authors pay careful attention to the possibility that inferior labor quality explains the lower productivity of labor in the east. One hypothesis is that workers' experience under communism is less valuable in a capitalist economy, so that in effect some human capital was lost by the transition. They confirm earlier findings by Alan Krueger and Jörn-Steffen Pischke that the return to experience fell between 1989 and 1990, but they also suggest that the experience gap may be less than it appears. Comparing compensation in east and west for workers with comparable "skill"—which they define in terms of a number of characteristics including age, education, and job tenure—they investigate how the east-west difference

changed between 1990 and 1999. To do this, they estimate the reward to skill from a wage equation run on a sample of westerners during the period 1984–89. They then use the estimated coefficients from this equation to predict what individual workers in the east would have received in the west in 1988. For 1990 a plot of these easterners' actual wages in the east against their predicted wages in the west shows a substantial gap at the mean skill level, with the gap growing as skill increases. The result is similar for both men and women. It indicates that the returns to skill were lower in the east than in the west and, if wage rates reflect productivity, suggests that the scarcity of capital in the east was particularly detrimental to skilled labor. By 1999 eastern wages had largely caught up to what would have been expected in the west in 1988, and the skill discount had disappeared. Indeed, for women it appears that skill is, if anything, more highly rewarded in the east.

The authors also consider a variety of other factors that may have been responsible for the productivity growth slowdown in the second half of the 1990s. They briefly discuss the potential role of shortages of managerial talent, entrepreneurship, and marketing skills and the possible impact of credit constraints. Another candidate is the slowdown in the growth of economic infrastructure in the east. At the start of the decade, eastern Germany's infrastructure was severely deficient. Immense outlays since reunification have narrowed that gap, but the pace of that narrowing has slowed with time, and the east's infrastructure is still regarded as inferior to the west's.

To explore the possibility that a slowdown in narrowing the infrastructure gap could help account for the observed pattern of productivity growth, the authors regress the annual Solow residuals, for both eastern and western states, taken from their previous analysis for 1992–99 on real infrastructure per capita (in some cases disaggregated by type), unemployment, and a set of variables characterizing the size of firms, along with eastern and western dummy variables interacted with unemployment. The equations also include fixed effects for each eastern and western state and separate time effects for east and west. The authors find a substantial and statistically significant effect for total real infrastructure and for infrastructure disaggregated by type. Their results indicate that a 10 percent decline in the growth of infrastructure equipment—not a large change relative to its growth rates in the decade—would reduce TFP growth by approximately 0.8 percentage point.

The same equations enable Burda and Hunt to examine the possibility that firms hoarded labor when output slowed after 1995, expecting the slowdown to be temporary. The potential importance of this phenomenon can be inferred from the estimated effect of unemployment on the TFP residual. For the west the coefficient on unemployment is negative as expected, and significant, but for the eastern states it is insignificant and has the wrong sign. The authors conclude that whatever hoarding took place in the east was overwhelmed by restructuring and other factors.

Many observers have pointed to the industrial structure of eastern Germany and the predominance of small firms as reasons for its low productivity. Between 1991 and 1995, manufacturing's share in employment dropped from over 25 percent to 15 percent, while construction's share grew from 10 percent to 17 percent. In 1990 eastern workers tended to be employed in larger firms than western workers; by 1999 the reverse was true. The authors look for clues to the importance of these factors in explaining aggregate wage differentials and the Solow residual. In their wage equations, differences in industrial composition and firm size have trivial effects on the size of the wage gap. They find that changes in firm size have a weakly significant effect on TFP growth: growth in the share of small firms in the east is estimated to have reduced TFP growth rates by 0.4 percentage point. They also explore whether migration, by removing the most productive individuals from the east, could be another factor depressing eastern productivity. However, disaggregating labor into four age groups and redoing the TFP calculations shows little difference from the aggregated results.

The extraordinary increase in eastern wages that quickly followed unification is undoubtedly one of the reasons for the rise in eastern unemployment. The authors associate this initial wage increase with several institutional developments. Western German labor unions were quickly established in the east, and a western German system of industrial relations was introduced. The unions were able to set relatively high wages in many industries, motivated in part by concerns for equity and the welfare of eastern workers, and the introduction of western welfare and social insurance raised reservation wages, reducing the cost of losing a job. The unions negotiated generous industrial agreements with employers' associations, some of which supported high wages as a way of retaining skilled workers who might otherwise emigrate. However, as it became clear that high wage rates were making their firms uncompetitive, an unprecedented

employer revolt ensued, with many firms leaving the employers' associations and abrogating their collectively bargained agreements. The importance of collective bargaining agreements in setting wage levels diminished steadily after 1993.

Although collective bargaining was thus responsible for the abrupt rise in wages, the authors reason that wages would have risen substantially, if more gradually, with integration even in the absence of powerful unions and generous welfare benefits. The convergence of wages was roughly in line with the convergence of productivity throughout the period, and productivity was bound to increase with the rapid transfer of capital and technology from the west. Furthermore, the option of emigrating to the west would have raised reservation wages in the east even in the absence of unions. As it happened, union membership in the east declined from nearly 50 percent of workers in 1993 to about 22 percent in 2000, compared with a drop in the west from 25 percent to 18 percent. Meanwhile the share of employees working for firms belonging to employers' associations declined from 76 percent in 1993 to 34 percent in 2000, by which time 29 percent of employees were being paid below union wages for their industry and region. There was little further convergence of eastern wages to western levels. But neither was there any important widening of the wage gap.

Most analysts believe that the unions raised wages generally to levels that were too high for full employment. A related question is whether they pursued equality at the expense of the relative wage changes needed to achieve an efficient reallocation of labor across sectors. Burda and Hunt report that inequality, as measured by the difference in log wages between the 90th percentile of the distribution and those at the 10th, in fact rose rapidly in the east and by 1999 had reached western levels. Wage equations estimated for east and west show that this growth in inequality is roughly equally divided between within-group and between-group increases (where groups are defined by age, education, tenure, and characteristics of the job, firm, or industry). The authors use similar equations to calculate how much of the growth in wages between 1990 and 1999 reflects changes in observable characteristics of workers, and how much reflects differences in the returns to these characteristics. They find that very little of the increase in wages is due to changes in observable characteristics. In earlier work Hunt found that low-paid workers had seen the sharpest fall in employment and the largest wage gains. Although this seemingly sup-

ports the hypothesis that the unions have caused unemployment by raising wages, particularly for the low skilled, the same pattern of employment losses occurred in the west, where wage gains for the low skilled were no higher than for others. This leads the authors to examine the experiences of the full range of skill groups. From the 1990 sample they create twenty skill groups of equal size, measuring skill as they did in their analysis of wage convergence. They then calculate the increase in wages and the change in employment for each skill group over the period. When the definition of skill takes into account employer characteristics such as firm size, high-skilled workers do substantially better than low-skilled workers. However, if these characteristics are ignored in defining skill, this verdict is, if anything, reversed. The authors regard these results as too weak to force a rejection of their strong prior beliefs that wage increases were not efficiently distributed across firms and industries in the east and have thus contributed to low employment.

Burda and Hunt believe the primary motivations for emigration prior to the parliamentary elections in 1990 were twofold—political reasons and family reunification—but that after that date economic reasons became dominant. To explore the importance of economic factors, the authors estimate an equation explaining gross migration flows between states by the wages and unemployment rates in the source and the destination state, allowing for time effects and fixed state effects and controlling for east-to-west migration flows relative to within-west flows. The results suggest that wage differentials are the most important economic factor in explaining migration and are more important at the start of the period than later. Unemployment in the destination state is important, but contrary to the usual presumption, unemployment in the source state is not significant. Disaggregating by age group, the authors find that wages are important for all ages, with the youngest workers most responsive. Unemployment in the source state has a substantial and significant effect on the migration of the oldest group, but no effect on the migration of the youngest. With the young accounting for a disproportionate share of migrants, this unresponsiveness presumably accounts for the insignificance of source-state unemployment in the aggregate results.

Burda and Hunt conclude on a note of guarded optimism. Looking to the future, they observe that the overarching question is whether eastern Germany will resume the process of convergence that had been so promis-

ing during the 1992–95 period, when wages, productivity, and consumption grew rapidly. The east is still dependent on transfers from the west, the labor market is still in disarray with unemployment extremely high, and TFP growth has fallen below western levels. But the authors regard the convergence of the wage structure to that of the west as a sign that wages are becoming more flexible and less of an impediment to the efficient distribution of labor and capital. Although it is hard to allocate the blame for the east's productivity shortfall among the wide range of contributing factors, they believe there are some ways in which public policy could help. Public infrastructure still lags behind the west in some areas, and the authors recommend continued investment in public goods. Despite increased wage flexibility, labor markets continue to need attention. However, Burda and Hunt believe that public training programs are overdone, and they are skeptical about public works jobs. At the same time, they argue, the German authorities should be more concerned with getting incentives right, for example limiting the ability of the unemployed to receive benefits after rejecting job offers. However, they maintain that the services offered to the unemployed through public labor offices are beneficial and should be expanded.

JAPAN'S STAGNATION SINCE THE start of the 1990s has moved it from the front of the growth pack among industrial economies to the rear. And with the recession deepening during 2001, its economic performance is getting worse, not better. Most analysts saw Japan's initial economic slump in the early 1990s as a reaction to the bursting of its real estate and stock market bubbles. The exposure of the banking sector to these asset price declines was a serious problem, but few if any observers expected that bad loans and insolvent banks would still hang over the economy a decade later. In the intervening years, a succession of Japanese governments have announced macroeconomic policy actions to restore growth. But except briefly in the mid-1990s, the economy has continued to stagnate or decline. By the end of the decade, already-large budget deficits and falling prices had led many to conclude that neither fiscal nor monetary stimulus could be effective, and cleansing the banking system of insolvent banks and bad loans had become the main policy prescription. In the second paper of this issue, Kenneth Kuttner and Adam Posen reexamine this long period of Japanese stagnation, which they and others call the Great Recession.

Their aim is not only to better understand Japan's situation, but also to use this experience to test conventional views about the effectiveness of fiscal and monetary policies.

Kuttner and Posen begin with a brief review of the decade, including an assessment of Japanese output growth relative to various estimates of its potential. Because Japan's growth rate changed so abruptly after the 1980s, there is little agreement on true potential output, and hence on the size of the gap between actual and potential at any point in time. The Organization for Economic Cooperation and Development (OECD) provides each year an estimate of Japan's potential; this estimate is based on projections of potential inputs that are then entered into an aggregate production function, together with an assumption about the rate of unemployment above which wages accelerate—the wage NAIRU. The authors also calculate an estimate of potential based on a latent-variable statistical methodology, which relies on two main assumptions: that actual output eventually reverts to potential, which is unobservable directly but evolves smoothly over time; and that the behavior of inflation is reliably linked to the output gap, which serves a purpose similar to that of the wage change relation in the OECD estimates. By both these measures of potential, Japan in early 2001 was operating 3 to 4 percent below its potential.

However, the authors reason, both these estimates may understate the growth of potential and thus the size of the output gap. They note that price and wage stickiness can make short-term inflation an unreliable indicator of the NAIRU or of the output gap in the two models. Furthermore, in the latent-variable methodology, with its assumption of mean reversion for actual output growth, any prolonged period during which the economy operates below potential will lead to an artificial downward shift in the estimate of potential. Moreover, the authors judge that the OECD estimates are based on projections of labor and capital inputs at potential that are likely to be too pessimistic: labor inputs because they ignore rising retirement ages and increased participation of women, capital inputs because they do not properly account for quality improvements in recent vintages. To underscore these possible shortcomings of the available estimates, they recalculate potential output combining actual labor inputs during the 1990s under the assumption that potential labor productivity growth since 1992 has continued at its average annual rate over 1980–92 of 2.5 percent. This calculation results in an 11 percent output gap in 1999. Even a slowdown to 1.5 percent annual productivity growth would

imply an 8 percent output gap by the end of the decade using this method. Although together these various estimates highlight the uncertainty in quantifying Japan's underperformance, its persistence is setting new records for an industrial economy in the postwar period, and by any reckoning the output gap has grown since the end of the 1990s.

Monetary policy is the main line of defense against an economic slump, and Japan's policymakers have been criticized for not using it more forcefully earlier in the 1990s. More recently, as prices have declined, the concern has grown that monetary policy can no longer be effective. The interest rate channel, through which monetary policy is usually thought to operate, is effectively closed, because short-term interest rates are already near zero. And there is a risk that increasingly negative inflation rates will raise real interest rates in the future. The authors thus explore whether monetary policy might be an effective stabilizer through other plausible channels. The main candidates would raise inflation or expectations of inflation, or would depreciate the currency, and they would presumably have to operate through monetary aggregates that the central bank can influence. This leads the authors to look for evidence of a connection between monetary aggregates and these potential policy targets.

The authors' statistical tests of the historical behavior of M2 and the main price indices in Japan show that they are difference stationary, meaning that their changes in response to a shock tend to be highly persistent. Further tests reveal M2, prices, and real output to be cointegrated, indicating that a connection might exist between M2 and prices that monetary policy could exploit. Estimated impulse responses based on a vector autoregression (VAR) of the three variables predict that persistent M2 shocks would produce persistent effects on prices. However, parallel tests using the monetary base—the aggregate that the central bank can affect most directly—reveal no historical cointegration with prices and give poor out-of-sample predictions of prices for the 1990s. The authors conclude that, even if faster M2 growth could be relied on to raise prices, the central bank would be unable to affect M2 if, as during the 1990s, banks are unable or unwilling to lend the additional reserves.

Altering the exchange rate of the yen would be another way of affecting the real economy. However, the authors note that the conventional route to depreciating the currency, by lowering interest rates, is not available to Japan. They present results of statistical tests showing that no other link from the monetary base to the exchange rate is important once interest

rates are accounted for, and they question whether, in the present environment, even the small residual effect that they find would be operative. This leaves them to consider the plausibility of monetary policy announcement effects. As Paul Krugman and others have argued, a credible commitment by the central bank to high inflation in the future could stimulate the economy in the present low-interest-rate environment. While conceding the logic of such arguments, the authors are skeptical that announcements by policymakers could have the needed effect on the inflationary expectations of businesses and consumers. And they provide evidence that recent policy announcements in Japan have not influenced bond rates. This leads them to conclude that monetary policy's ultimate role in the present situation may, by necessity, be limited to accommodating an expansionary fiscal policy.

Hence the authors turn to an examination of countercyclical fiscal policy, whose effectiveness has also been questioned in the present environment. Skeptics argue that Japan has taken conventional, expansionary budget measures and that they have failed. One problem with evaluating this claim is that budget data for Japan, where responsibilities are shared between the central and local governments, are not as transparent as are those for the United States. Generally accepted statistics from Japan's Economic and Social Research Institute indicate that a sizable surplus at the start of the decade gave way to deficits beginning in 1993, which grew almost without interruption in each subsequent year. By 1999 the deficit on this measure was ¥35 trillion, or 6.8 percent of GDP, and central government debt was roughly equal to annual GDP. However, the authors point out that these budget outcomes are themselves importantly affected by the stagnation of the economy. Revenue and transfer expenditure depend on the level of economic activity, and rising deficits during a period of stagnation can result from this reaction of the budget as an automatic stabilizer, just as rising deficits during brief recessions do. The conventional way of dealing with this issue is to calculate what the budget surplus or deficit would have been if the economy were operating at potential. Changes over time in this measure can be interpreted as changes in the thrust of fiscal policy, since they are uncontaminated by the endogenous reactions of the budget to deviations from the potential output path.

Although these considerations make it clear that Japan's actual deficits over the 1990s are not a useful basis for evaluating fiscal policy, in light of the great uncertainty about the potential output path of the Japanese

economy, Kuttner and Posen are reluctant to offer estimates of the budget surplus at potential. Instead they estimate fiscal impacts using a structural VAR model in which real tax revenue, real expenditure, and real GDP are mutually endogenous variables subject to shocks. The model, adapted from work by Olivier Blanchard and Roberto Perotti, is estimated using annual data for fiscal years 1976–99 and combining the accounts of the central and local governments. It allows for contemporaneous effects of GDP on receipts but not expenditure, and for no contemporaneous effects between receipts and expenditure except for an effect of tax shocks on spending. All other interactions are lagged one year. Although such restrictions could influence the estimated parameters, the authors believe they correspond to the institutional setup for fiscal policy in Japan. The estimation procedure also allows for a linear trend and a trend interacting with a post-1990 dummy variable, and to permit the other parameters to be identified, it fixes the elasticity of revenue to GDP at 1.25, a value taken from other studies.

The estimated equations show that current expenditure and lagged revenue have significant effects on GDP, with the expected signs, indicating that fiscal policy is effective in the conventional way. Calculating the dynamic effects of shocks to taxes or expenditure over a four-year horizon, the authors again find support for the effectiveness of fiscal policy. A ¥100 tax cut raises GDP by a cumulative ¥484, and a spending shock of the same size raises GDP by ¥353. Such estimates do not, however, measure the conventionally understood impact of a one-time policy shock. They take as endogenous any changes in taxes and spending after an initial shock, thus embodying in the estimation any historical tendency of policymakers to either augment or offset any initial policy shock as well as any effects of the automatic stabilizers. The authors suggest that the tendency for taxes to rise in the years after a rise in expenditure probably accounts for the observed difference in the estimated dynamic effects of taxes and spending.

The authors augment this general analysis by addressing specifically whether fiscal policy in Japan has become ineffective, as some observers argue. They estimate a time-series regression of a form first suggested by Michael Hutchison, regressing private saving (adjusted for social security) as a percentage of GDP against a number of alternative fiscal variables, also expressed as percentages of GDP, and the old-age dependency ratio, which itself might have affected saving in this period. Using the overall

budget balance as the fiscal variable, they find that the demographic variable has a positive (although insignificant) coefficient, a result that is contrary to what the life-cycle hypothesis would predict when, as in Japan, population aging is due to a declining birth rate. The effect of the budget balance, although significant and negative, is small, implying that only 12 percent of a change in the budget balance is offset by reduced private saving. The authors point out that this is very far from the 100 percent offset predicted by Ricardian equivalence. When tax revenue and government expenditure are included in the equation in place of the overall budget balance, the authors estimate that 32 percent of a change in revenue is offset by private saving, whereas expenditure, although insignificant, has the wrong sign, with private saving augmenting rather than offsetting the effect of the budget balance. Neither the social security balance nor interest expenditure is significant when added to this equation. The authors conclude that Japan's experience offers little support for the Ricardian notion that private saving offsets public dissaving. They further observe that, in contrast to some countries where fiscal contraction appears to have been expansionary because they seemed to be approaching binding borrowing constraints, Japan has had low interest rates and little or no foreign-held or foreign currency–denominated debt.

Kuttner and Posen turn finally to a discussion of Japan's banking crisis and the effects it may be having on the real economy. They sketch the evolution of Japan's postwar financial system and review how the crisis emerged and why the government's unwillingness to deal with it decisively at the start only made matters worse as the decade progressed. But their main purpose is to examine whether the banking crisis itself has contributed to the stagnation of the real economy. Bank lending will fall in response to weakness in real activity that reduces loan demand. But the causality can also run the other way, with a reduction in the willingness or ability of banks to supply loans contributing to weakness in aggregate demand and real activity. Ben Bernanke has identified bank closures as a factor contributing to the depression of the 1930s. With Cara Lown, he has shown similar effects from the thrift industry crisis in the United States in the late 1980s. However, Kuttner and Posen note that analysis of Japan in the 1990s has generally concluded that lending was constrained by demand rather than supply.

Kuttner and Posen investigate this issue with a set of regressions relating, alternatively, bank lending, business investment, and the *tankan* sur-

vey of bank lending attitudes to macroeconomic variables that could be expected to affect loan demand and one of two proxies intended to capture the degree of stress in the banking system: changes in the real price of commercial real estate and changes in the real stock price of banks. They also include two dummy variables to capture changes in the competitive environment and in the regulatory environment. The first covers the first half of 1998 through the first half of 1999, when Japan's banks had to pay a premium in international markets for overnight money and when private savers moved money from banks to postal savings accounts. The second covers the first half of 1998 through the second half of 2000, a period of heightened regulatory scrutiny. The regressions using real estate prices to proxy for bank stress start in 1976, and those using stock prices start in 1983.

From these regressions, the authors conclude that bank stress produced significant reductions in the supply of loans from banks and, through this channel, contributed to weakness in the real economy. The two bank stress proxies are always significant, helping explain real loan growth in both city (nationwide) banks and regional banks, *tankan* lending attitudes, and real investment. The authors calculate that the decline in real estate prices between 1992 and 1995 had an effect on bank lending equivalent to the regressions' estimate of the effect from a 3-percentage-point rise in interest rates. They also calculate that the sharp decline in stock prices in the first half of 1992 caused a 3 percent decline in lending. The dummy for the period of increased competition is also significant, whereas the dummy for intensified regulatory scrutiny is not; this, the authors suggest, may show that scrutiny has not intensified much after all.

In closing, Kuttner and Posen discuss what policymakers can learn from Japan's experience. Their own analysis leads them to conclude that the conventional countercyclical tools of fiscal and monetary policy work. Tax cuts and spending increases are expansionary, and Ricardian equivalence effects are not large enough to be relevant to stabilization policy. They believe that, in the first half of the decade, monetary policy focused too much on inflation, and that subsequent developments showed that a liquidity trap is a distinct threat to the conventional and normally reliable interest rate channel for stabilization. A general lesson they take from this experience is that monetary policy should not fear accelerating inflation when inflation is low. They observe that financial fragility can affect the real economy, and that avoiding needed bank closures and forcing banks to roll over loans, as Japan's authorities have done, does not avoid these real

effects and may only increase the eventual cost of banking reform. The authors also call attention to two important issues raised by their review of Japan's experience but not directly addressed in their analysis. One is whether Japan would have been better off if it had been more integrated with international capital markets. They suggest that stabilization tools might then have been less effective, but that the discipline of international capital flows might have kept them from getting into so much trouble in the first place. The other is why Japan's political system was so unresponsive to the structural and stabilization problems that emerged. They leave this question as an important topic for future exploration.

SOCIAL WELFARE POLICIES ARE WIDELY known to be more generous in Europe than in the United States. Income taxes in Europe are typically more progressive, transfers more generous, social safety nets stronger and more encompassing, and regulations to protect workers and the poor more extensive. There is, of course, variation among the European countries, but even the United Kingdom, the least generous, provides more social welfare than the United States. The implications of these differences for economic performance and for the well-being of the typical citizen have generated considerable research. In the third paper of this issue, Alberto Alesina, Edward Glaeser, and Bruce Sacerdote explore a different question: Why is Europe so different from the United States in this central dimension of public policy?

The authors first show that the differences are not of recent origin. Although the expansion of public sectors on both sides of the Atlantic began in the late nineteenth century, from the outset the European countries led the way. In 1870 U.S. government expenditure on subsidies and transfers was 0.3 percent of GDP, compared with 0.9 percent in Europe. Those shares were 2.1 percent and 6.8 percent, respectively, in 1937; 5.0 percent and 11.5 percent in 1960, when a significant expansion of welfare programs began; and 11.0 percent and 21.0 percent in 1998. This long history of persistent and growing differences leads the authors to exclude explanations that are specific to certain periods or events.

The authors also document differences over a range of social welfare categories. Examining Germany, Sweden, and the United States in some detail, they show that benefits for families with children, for health care, for sickness and accidental injury, for disability, and for poverty relief are all less generous and of shorter duration in the United States. In all the

European countries examined, minimum wages replace a larger fraction of average wages than in the United States, labor market regulations are more protective of workers, and unemployment benefits last longer and (except in the United Kingdom) are more generous. The fact that this greater European generosity prevails across a range of social goals suggests that there are broad underlying reasons for the differences.

The authors provide an analytical discussion of the way individuals' preferences, efficiency cost, income mobility and the uncertainty of future income, and characteristics of the political system affect the political process that determines the amount of income redistribution. Guided by this discussion, the authors turn to an empirical investigation that attempts to account for the observed differences in redistribution. They first look at economic explanations but find ambiguous results. Greater before-tax inequality might be expected to lead to greater redistribution. That would not appear to help explain U.S.-Europe differences, however, because Gini coefficients reveal noticeably greater before-tax inequality in the United States than in Europe. On the other hand, the authors note that redistribution in Europe could take place in ways that affect the observed before-tax income distributions, for example through higher minimum wages. Another factor, a lower perceived efficiency cost of redistribution, might also be expected to matter. But the authors find little evidence that taxes are less distortionary in Europe than in the United States. They note that differences in income mobility—the ease with which a household moves up or down the income distribution—could also help explain differences in redistribution, although the sign of the effect depends on what drives income mobility. If voters near the median of the income distribution expect above-average income growth in the future, they are likely to favor less redistribution, whereas if they face a greater risk of exogenous shocks to their income, they will likely favor more. The authors review various earlier studies of the mobility and variability of income but find no clear verdict. Thus none of their evidence on economic explanations helps account for the observed differences in redistribution.

Alesina, Glaeser, and Sacerdote next examine possible political explanations for Europe's greater welfare generosity. They first focus on the degree to which political systems are based on proportional representation. This form of representation gives greater voice to minorities, who, the authors hypothesize, tend to favor transfer programs, than does majoritarian representation, in which the preferences of the median voter

dominate. Two data sets are available for this analysis, each from a previous study by other authors. The first uses data for general government for the countries of the OECD and Latin America; it measures proportionality by a continuous variable that attempts to capture the share of electoral votes that guarantees a party a parliamentary seat in a typical electoral district. The larger the needed share, the further the system is from proportional representation, since it is then more difficult for small parties to win a seat. In cross-country regressions explaining transfers as a share of GDP in just the OECD countries, and including GDP per capita and the percentage of the population over age sixty-five as control variables, the proportionality variable (which tends to be higher in Europe) is highly significant.

These OECD results do not seem to apply across a broader set of countries, however. When the Latin American countries are included in the sample, and a dummy variable for Latin America is added to the regression, the proportionality variable remains significant, but its coefficient is cut in half. The reason for the smaller coefficient is revealed in a scatterplot, which shows that the effect of proportionality in the Latin American countries has the opposite sign from that in the OECD countries.

Turning to their second data set, which includes many more countries, the authors are forced to change some of the variables in their analysis. The dependent variable is now social spending as defined by the International Monetary Fund, and two new political variables are introduced: the first simply indicates whether a country has a presidential regime, and the second assumes the value of one if a country has a majoritarian electoral system and zero otherwise. These equations also include additional regional dummies. Neither political variable is significant in this regression. The authors conclude that proportionality is significantly associated with redistribution across the OECD countries, but not across countries from the developing world.

The authors recognize that the regressions just described capture only one of many political and institutional features that distinguish the United States from Europe. Furthermore, they acknowledge that these features may themselves be endogenous, reflecting more basic preferences of the populations in either recent or distant history, and that these basic preferences may themselves be closely related to a taste or a distaste for redistribution. In this spirit they discuss historical experiences and concerns that

may plausibly have shaped the political arrangements that affect redistributive policies today.

Following their analysis of political institutions, Alesina, Glaeser, and Sacerdote turn to a wide-ranging exploration of theories about behavior, seeking to answer why the median voter in Europe might be more disposed than the median voter in the United States to support policies to help the poor or to favor redistribution more generally. Their discussion centers on evidence about racial prejudice and the behavioral concept of reciprocal altruism. Numerous studies identify racial discrimination as important in many settings, and the authors judge it to be more important in the United States than religious differences, which loom large elsewhere in the world. In the context of welfare, the concept of reciprocal altruism predicts that nonrecipients will oppose welfare spending if they believe that recipients are taking advantage of the system. The authors suggest that such beliefs may, in turn, be related to the racial makeup of the recipient population. They cite a number of studies whose results are consistent with the idea that race is important in shaping people's attitudes toward others. From such evidence they infer that the much greater racial heterogeneity of the United States, together with the disproportionate share of minorities among its poor, may be an important reason why the political taste for redistribution is lower here than in Europe.

To examine these ideas more formally, the authors conduct a number of regression analyses aimed at establishing the importance of various factors in shaping people's attitudes toward redistribution. They first report on cross-country regressions that use racial fractionalization, defined as the probability that two randomly drawn individuals in the same country are of different races, to explain social spending as a share of GDP. The data set is that used in the analysis of majoritarian regimes described above, covering a large number of countries worldwide. The majoritarian regime dummy, regional dummies, GDP per capita, and the percentage of the population that is of working age are additional explanatory variables. The authors find that racial fractionalization is significant and has the expected negative sign, whereas the majoritarian regime variable is also correctly signed but not quite significant. When they substitute ethnolinguistic fractionalization for racial fractionalization, it is not significant.

Using U.S. data for 1972–98 from the General Social Survey, the authors next investigate how individual Americans' views on whether

welfare spending should be increased correlate with various personal characteristics. In one regression they find that the single largest coefficient is on the race of the respondent, with African Americans 23 percent more likely than others to favor more welfare spending. They also find a significant negative coefficient on the respondent's income and a U-shaped pattern for educational attainment, with those with the least education (high school or less) and those with the most (postgraduate education) voicing the strongest support for more spending. Support for more welfare is also importantly related to living in a large city, whereas gender, marital status, and number of children are less important correlates. When the analysis is restricted to whites, three new variables, each introduced separately, are revealing. Support for welfare is negatively related to the proportion of blacks living in the same state (although the effect is not significant), significantly negatively related to the belief that blacks are lazy, and significantly positively related to whether the respondent has had a black person over to dinner within the past few years. The authors pursue the role of race further in a different set of regressions, this time relating the generosity of a state's Aid to Families with Dependent Children (AFDC) program in 1990, before the 1995 welfare reform, to the share of its population that is black. In a regression that also includes states' median income, they find that a 1-percentage-point increase in the share of the population that is black reduces a state's maximum AFDC payment by $7 a month. They see this set of results as supporting the idea that racial heterogeneity is one of the most important reasons why the welfare state in America is so much smaller than in Europe.

The authors next try to relate welfare differences to the concept of reciprocal altruism, briefly summarized as the idea that people respond in kind to the way others act toward them. An application of this idea would be that people who believe welfare recipients are taking advantage of them will tend to disfavor income support programs. Citing as a classic example Ronald Reagan's apocryphal welfare queen living high on taxpayer dollars, the authors observe that antiwelfare forces often frame the political discussion along such lines. But for their purpose the more interesting question is why this discussion is so different in the United States than in Europe. Opinions about poverty, they point out, are very different in the two places.

What forces might account for such different opinions? The authors offer a number of observations that they feel provide clues. There is a

strong positive correlation between earnings and hours worked in the United States, but a much weaker pattern in European countries. The American ethos for hard work may reflect Puritan beliefs that still influence attitudes today. Negative attitudes toward welfare recipients may also stem from the greater social isolation of the poor in the United States. In a different vein, the authors suggest that Americans may be more comfortable about punishment in general—Americans overwhelmingly support the death penalty and believe that the courts are too lenient in punishing criminals—and such an outlook may carry over to punishing welfare beneficiaries.

Again using data from the General Social Survey, the authors test these propositions among U.S. respondents, at the same time examining the possible role of occupational mobility, which they measure as the mean difference between the occupational prestige of a respondent and his or her father. Here the idea is that people who have bettered themselves are likely to feel that the poor could do the same if they tried. The authors regress support for welfare on occupational mobility, support for capital punishment, church attendance, and whether the respondent is Protestant, along with the range of personal characteristics used in the regressions on race described above. All the new variables are negatively related to support for welfare, and the role of personal characteristics is as before.

Turning to cross-country data from the World Values Survey, the authors find important differences in beliefs about the poor between Europeans and Americans. By wide margins, relatively more Europeans believe that the poor are trapped in poverty, that luck determines income, and that the poor are not lazy. They also examine whether these beliefs, and their opposites, are related to self-identification as being on the left wing of the political spectrum, which the authors take as a proxy for supporting welfare. However, these results do not lend strong support to the idea that left-wing orientation is consistently associated with these beliefs. Nor do other cross-country regressions, including attempts to relate social spending to the belief that luck determines income, and attempts to explain leftist political orientation, uncover significant relations.

The authors' answer to why redistribution is so much greater in Europe than the United States thus centers on personal attitudes and on political institutions. They find little role for economic explanations. Their conclusion emphasizes that the majority of Americans believe that redistribution favors racial minorities, and that they live in a fair and open society, so that

the failure of any individual is his or her own fault. And they suggest that the American political system, which is geared toward preventing redistribution, is likely to be endogenous, reflecting these basic beliefs.

OVER THE LAST CENTURY THE average annual return on stocks has exceeded that on short-term government bonds by more than 5 percentage points. Economists have long been puzzled by the magnitude of this equity premium. Although equities are much riskier than bonds, such a reward for risk taking is much larger than that predicted by most empirical studies based on the standard models of consumer saving and portfolio choice. There have been numerous attempts to explain the puzzle, invoking nonstandard consumer preferences on the part of consumers or costs and limitations on investment allocations, but none have been notably successful. In the fourth paper of this issue, Jonathan Parker provides a possible explanation that stays within the framework of the standard model.

According to the standard model, risk premiums on stocks should reflect a negative correlation between equity returns and the marginal utility of income. Since we expect consumption to be positively correlated with equity returns, marginal utility should be low when stock returns are unusually high, and vice versa. Hence positive excess returns add less to utility than negative excess returns of the same size subtract, so that, for any given expected return, equities have a lower value than risk-free assets. But as measured empirically, the contemporaneous, quarter-by-quarter covariation of consumption with the stock market is low. As a consequence, it requires an enormous, and implausible, sensitivity of marginal utility to consumption to rationalize the historical risk premium. To address this puzzle, Parker proposes two variations on the empirical application of the model: the first is a focus on "medium-term risk," or the correlation of consumption with lagged rather than contemporaneous equity returns; the second allows for the possibility that the covariation of market returns with the consumption of those who actually hold stocks is much greater than their covariation with aggregate consumption per capita.

Parker begins by reviewing the canonical model for explaining the consumption, saving, and portfolio decisions of households and showing why it implies an incredibly high degree of risk aversion as conventionally estimated. According to the model, households seek to maximize the present discounted value of utility from consumption, where utility is additively

separable and the period-by-period utility function is assumed to be of the constant-relative-risk-aversion form. An optimizing household must be indifferent between investing in the risk-free asset and investing in equities. Assuming that returns on equity and consumption growth are conditionally log-normally distributed, it is straightforward to show that this implies a degree of risk aversion, for a given expected value and variance of the excess return, that is inversely proportional to the covariance between the consumption growth rate and the excess return. In principle, these variances and covariance should be conditional on information available at the time of an investment decision, but Parker shows that empirically it makes little difference whether one uses conditional or unconditional estimates of variances and covariances.

In U.S. quarterly data from 1959 to 2000, the average annual equity premium is about 5.3 percentage points, the unconditional standard deviation of excess returns is about 16 percent, and the unconditional covariance is 0.00017. This implies a degree of risk aversion for the typical household of roughly 380. Parker observes that economists would find a value of about 4 plausible, and anything above 10 implausible. A risk aversion of 10 would imply that a household would give up 19 percent of its consumption with certainty in order to avoid the risk in a 50-50 chance of increasing or reducing its consumption by 25 percent. A value of 380 would imply that the household would give up more than 24 percent of its consumption to avoid that same risk. Parker also illustrates the puzzle another way: with a risk aversion of 4, a household facing risk as measured by the contemporaneous covariation should be indifferent between holding stocks and holding bonds only if the equity premium is less than 0.1 percentage point, not the 5.3-percentage-point historical average.

Parker then turns to a discussion of the rationale for measuring consumption risk as the medium-term risk, or the covariation of the one-period excess return with consumption growth over a horizon of S quarters. He shows analytically that medium-term risk is approximately correct even if risk calculated from contemporaneous covariation as in the textbook model is exactly correct, whereas medium-term risk is superior to the textbook model under a wide variety of possible deviations from the model. In principle, an innovation in equity returns could lead to a change in future risk-free rates, leading in turn to a significant revision in planned consumption over the following periods. In that case Parker's measures of medium-term risk would be inappropriate. If, however, future

innovations are the only reason for revising future consumption plans, medium-term risk properly measures consumption risk. He shows that, empirically, the error in making this assumption is insignificant.

On the other hand, there are many plausible reasons why the calculation based on the textbook model would overestimate the degree of risk aversion, and the medium-term risk measures would be more appropriate. Parker describes three types of reasons. First, suppose market returns are correlated with transitory factors that would affect the marginal utility of next period's consumption. Then changes in the marginal utility of consumption are not well measured by contemporaneous consumption growth. As the horizon lengthens, the correlation of such factors with the current equity return is likely to fall, making Parker's measure of medium-term risk more accurate. Constraints on consumers' access to information, costs of calculation, or costs of adjusting consumption are examples of a second type of reason, all implying that consumption, and hence marginal utility, adjust slowly to the new optimal level following a shock to equity returns. In a recent paper, Xavier Gabaix and David Laibson present a fully worked out example of this type in which the costs of monitoring lead households to check their portfolios and adjust their consumption infrequently. Finally, reasons of the third type involve measurement errors such that aggregate consumption data measure consumption responses with a delay, even if the true response is instantaneous.

For his empirical work, Parker considers consumption horizons of up to eleven quarters and uses two alternative consumption measures, each with its advantages. The first is real consumption per capita on nondurable goods and services less expenditure on education services, medical care services, housing services, personal business services, and footwear. This measure, which he calls "flow consumption," seems to be a good approximation to the theoretical concept of consumption, removing from the usual measure expenditures that have a substantial element of investment (such as education) or reflect regrettable necessities (such as medical care) that do not affect marginal utility in the usual way. Parker's second measure is total consumption expenditure less expenditure on education, medical care, and personal expenses. He recognizes that this measure, which does not exclude durable goods purchases, is unorthodox, since utility comes from the service flow from durables, not from expenditure on them. However, he argues, over long horizons the new level of expenditure on durables is proportional to the new level of the stock, and hence to the new

level of durable services. Since increased consumption of the services from durable luxury goods may be an important part of the response to unexpected high returns, this consumption measure may be appropriate when calculating medium-term risk.

Parker first calculates unconditional covariances of excess equity returns and consumption growth using his two measures of consumption. As previously noted, the contemporaneous covariance of returns and flow consumption is very low, indeed statistically insignificantly different from zero. Given the variance of equity returns, this covariance implies flow consumption growth of only 0.06 percent above average when equity returns are 10 percent above normal. The contemporaneous covariation of total consumption expenditure with returns is even lower. In both cases the covariance jumps dramatically when the horizon is extended to one quarter ahead; it then increases gradually as the horizon stretches further, so that at a consumption horizon of six quarters the covariance exceeds the contemporaneous covariance by an order of magnitude or more. This implies that estimates of risk aversion using these longer horizons are approximately one-tenth those implied by the contemporaneous measures of risk.

The estimates so far make no attempt to utilize current information to improve on unconditional forecasts of either excess returns or consumption. Parker observes that, empirically, the stochastic process of equity returns shows that, after a series of particularly high returns, subsequent returns are on average lower, and vice versa. Eliminating the predictable component of returns may give different estimates of risk aversion and narrow the confidence intervals on those estimates, which are quite large when using unconditional covariances, even when uncertainty in the premium itself is ignored. To investigate this possibility, Parker estimates a three-variable vector autoregression (VAR) relating excess returns, the logarithm of consumption, and the dividend-price ratio, which historically has been a relatively good predictor of future returns. The system not only allows conditional forecasts of future returns, but also provides a clear picture of the response of consumption to innovations. Each variable is included with four lags. Excess returns are ordered first so that the impulse responses measure the impact on consumption of an innovation to returns. Parker notes that innovations to returns should not be taken as structural, but rather as an amalgam of structural shocks to the economy such as news about labor income or future returns. Hence the estimated impulse

response does not measure the marginal propensity to consume, even though it is relevant to the measurement of medium-term risk. This system is estimated for both consumption measures.

The impulse responses are interesting on their own. For example, for either consumption measure, most of the consumption response has taken place by the end of five quarters, but even at that horizon consumption has increased less than 0.1 percent for a 1 percent innovation in excess returns. Using the estimates from the VARs, Parker computes medium-term risk for the various horizons and the implied risk aversion coefficients. The latter are in the range of 30 to 40 at a horizon of a year, similar to those for the unconditional estimates. The conditional estimates are more precisely estimated, however, with standard errors roughly two-thirds as large as those for the unconditional estimates.

Using data from the entire sample period to estimate the response of consumption to excess returns takes advantage of information not available to households investing in real time. To see whether this artificial element in the VARs is important to his conclusions, Parker performs a similar exercise using the same predicting variables but running a series of rolling regressions, so that the information set used to predict returns includes only data available to agents at the time. The results are much the same for horizons up to eight quarters. Beyond that they diverge but are highly uncertain.

Parker concludes that although the medium-term risk to consumption is much greater than that calculated from the contemporaneous covariation of returns and consumption growth, it is still much too low to resolve the equity premium puzzle. In an attempt to narrow the gap further, he invokes the other feature of the stock market that he regards as potentially important, namely, the fact that only a fraction of households own equities.

Because so many households do not hold equities, the covariation of fluctuations of aggregate consumption with stock market returns may be a poor indicator of the risks faced by households that do. It would not be surprising, for example, to find that the consumption of equity holders is more dependent on equity returns than that of non–equity holders, indicating correspondingly higher risk for the former. To investigate this issue, Parker uses data from the Consumer Expenditure Survey (CEX) to construct a series measuring the consumption growth of equity-owning households. He believes this survey contains the best household-level data on consumption over time in the United States; it also includes not only

demographic information but also the estimated market value of holdings of stocks, bonds, and other financial assets. Using these household data, Parker constructs flow consumption and total consumption expenditure per effective person in the household. After dropping outliers, Parker averages the growth rates of consumption per effective person to arrive at a measure of the growth of consumption for stockholding households.

There are advantages and disadvantages to the use of CEX consumption data rather than the aggregate data from the National Income and Product Accounts (NIPA). Parker notes that the CEX consumption series has significantly more measurement error than the aggregate series, but if the errors are proportional to consumption and have a zero mean conditional on true consumption, the calculation of the medium-term risk based on that series will still be unbiased. The CEX data also cover a shorter period than the aggregate data, so that a greater fraction of the data come from the recent period of unusually high excess returns in the stock market, and, possibly, of higher covariances of returns with consumption. To provide a baseline for comparison, therefore, Parker redoes the aggregate analysis for the same period covered by the CEX data, 1979–98. He finds that even though the medium-term risk for this period is considerably higher, if the higher average returns are taken to be expected, the corresponding risk aversion coefficients are typically roughly double those estimated for the entire sample.

Is the medium-term risk calculated for equity owners from the CEX data greater than the risk calculated from the aggregate data for the same subperiod? Parker finds that, for both measures of consumption, the answer is yes. The contemporaneous covariances are at least five times larger than those estimated from the aggregate data, and as with the aggregate data, the covariances rise significantly as one extends the horizon. For the CEX data the medium-term consumption risk is two to four times greater than the contemporaneous risk, depending on the consumption measure used.

Whether investors expected the high rates of return and low covariances experienced in the 1979–98 period is crucial to these estimates. As mentioned above, with the aggregate NIPA data, the estimates of risk aversion for this subperiod were roughly double the estimates for the entire period. To allow for the possibility that the CEX estimates are unduly driven by the atypical experience of these years, Parker calculates the ratio of risk aversion estimated from the aggregate data for the CEX subperiod

to that estimated for the whole period, and then divides his CEX-based risk aversion estimates by this ratio. The resulting estimates of medium-term risk aversion range from 5 to 10 for flow consumption and from 3 to 8 for total consumption, well below estimates using aggregate data for the whole period.

The CEX data on stockholders differ from the aggregate NIPA consumption data in a number of ways. As Parker notes, one difference is that the CEX data on consumption growth are aggregated from individual households, whereas the NIPA consumption growth data not only confound the distribution of consumption with movements in the typical household's consumption, but do not account for the fact that some households die, immigrate, emigrate, or are "born" into the sample between periods. A second difference is that the NIPA consumption data include the spending of nonprofit organizations. A third is that the CEX covers a smaller, nonrandom sample and excludes rural and military households and students living in dormitories.

To explore how important these differences are, Parker constructs a pseudo-NIPA consumption series for all households from the CEX data by mimicking the NIPA method of aggregation. Estimates of the consumption risk of equity using this pseudo-NIPA series are lower than those reported above for stockholders, but higher than those using actual NIPA data; the degrees of risk aversion calculated from the pseudo-NIPA data are roughly half the NIPA-based estimates. Parker suggests that these calculations give a reasonable idea of the importance of the remaining differences, which arise from defining and aggregating consumption differently and from restricting the sample to stockholders.

To get an indication of the importance of each of these differences, Parker calculates the medium-term risk for all households in the CEX, defined and aggregated in the same manner as for stockholders. Calculations are done for both conditional and unconditional covariances. The medium-term risk for stockholders varies with the horizon and averages roughly a third larger than that for all households. By comparison, using the CEX definition and aggregation increases medium-term risk by roughly two-thirds.

Parker separately examines older households and relatively rich households, two subpopulations that might be expected to be more dependent on equities as a source of income and that might therefore be subject to greater consumption risk from equity than either of the populations con-

sidered thus far. He finds that older households have only slightly more medium-term risk than the typical household, possibly reflecting the fact that the elderly have substantial amounts of Social Security wealth, whose returns are safe and uncorrelated with the market. Parker finds no evidence that wealthy households, defined as those holding more than $25,000 in 1982–84 dollars in securities (corresponding to the top 36 percent of stockholding households), bear no more medium-term risk than the typical household owning stocks.

The finding that the medium-term risk of those actually holding equities is sufficient to resolve the equity premium puzzle does not resolve other problems. Most notably, as Parker acknowledges, and as stressed by Paul Willen in his discussion, it does not explain why many households hold no equity at all. But Parker's results do suggest that it is unnecessary to follow some recent research in focusing on significantly different models of consumption behavior or nonstandard utility functions to explain risk premiums. Rather, the results point to the need to adopt models of slow adjustment of consumption to explain not only the equity premium but asset pricing more generally. The success of the model also raises the tantalizing questions of whether equity risk has been reduced by the improved risk sharing that has taken place in recent years as participation in the market has broadened, and whether that broadening has permanently lowered the premium required for holding stocks.