

**COMPARATIVE MACROECONOMIC PERFORMANCE AND ALTERNATIVES TO
THE NATURAL RATE HYPOTHESIS***

J. Bradford De Long

Harvard University and NBER

Lawrence H. Summers

Harvard University and NBER

September 1988

PRELIMINARY

*We would like to thank Robert Barro, Robert Barsky, Olivier Blanchard, Bruce Lehmann, Greg Mankiw, George Perry, Christina Romer, Andrei Shleifer, and participants in seminars at Princeton and Northwestern for helpful discussions.

The natural rate hypothesis and its implication that demand management policies cannot affect the economy's average level of unemployment or output over long periods of time has come to be widely accepted even by Keynesian economists. This view is enshrined in standard textbooks: Hall and Taylor, for example, "stress... that macro policy cannot influence the average rate of unemployment. It can only influence the fluctuations of unemployment around the natural rate."¹

This paper raises questions about the validity of the natural rate hypothesis. It argues that demand management policies can and do affect not just the variance, but also the mean of output and unemployment. As a way of comparing intervals in which different demand management policies stabilized the economy to different degrees, we return to the much-discussed comparison of pre- and post-WWII macroeconomic performance in the U.S. and other industrial nations. Previous explorations of macroeconomic performance in historical perspective have focused on the volatility of output, either about trends or of changes.² But this is not the relevant measure if, as Keynes and the early Keynesians believed, successful macroeconomic policies fill in troughs without shaving off peaks.

Even using the pre-war GNP series created by Christina Romer to correct for excess cyclical variability in the standard Kuznets-Kendrick-Gallman series,³ we find evidence of considerable

¹Robert Hall and John Taylor, *Macroeconomics*, 2nd ed. (W.W. Norton, 1988), p. 486.

²See, for example, Arthur F. Burns, "Progress Toward Economic Stability," *American Economic Review*, 50 (March, 1960), pp. 1-19; J. Bradford De Long and Lawrence H. Summers, "The Changing Cyclical Variability of Economic Activity in the United States," in Robert J. Gordon, ed., *The American Business Cycle: Continuity and Change*, (University of Chicago for NBER, 1986), pp. 679-733; Charles Schultze, *Other Times, Other Places*, (Brookings, 1986); Martin Baily, *Stabilization Policy and Private Economic Behavior*, 1978 (Spring 1978), pp. 11-60; and Steven Sheffrin, "Have Economic Fluctuations been Dampened? A Look at Evidence Outside the U.S." *Journal of Monetary Economics*, 21 (January 1988), pp. 73-84.

³Christina Romer, "Spurious Volatility in Historical Unemployment Data," *Journal of Political Economy*, 94 (February 1986), pp. 1-37; "Is the Stabilization of the Postwar Economy a Figment of the Data?" *American Economic Review*, 76 (June 1986), pp. 314-34; "Gross National Product, 1909-1928: Existing Estimates, New Estimates, and New Interpretations of World War I and Its Aftermath," (Princeton, 1987); "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product," *Journal of Political Economy* (forthcoming 1989). Romer's work has sparked a debate that includes contributions from David

improvement in macroeconomic performance in the post-WWII period. The average “gap” between real aggregate demand and the potential supply of the economy has been reduced relative to potential by an amount which would now run at almost \$50 billion a year in the post-WWII as compared to the pre-Depression period. This improvement in the performance of the economy relative to potential is conventionally, and we think correctly, attributed to the more stable financial system, automatic stabilizers, and possibly improved discretionary macroeconomic policies that together make up the post-WWII Keynesian institutional order.

We first stress that the economy's average level of unemployment is likely to be inefficiently high; an increase in the average level of output is likely to be desirable whether or not the operation of the natural rate property makes it infeasible. We go on to review recent theoretical developments that suggest alternatives to the natural rate hypothesis. Recent years have seen the emergence of a variety of theoretical models which exhibit multiple equilibria associated with different levels of production. These models, which capture many of the ideas discussed in *The General Theory*,⁴ rely on mechanisms like credit failures, low-level demand traps, and asymmetries in price adjustment to generate economies with multiple equilibria at which the forces pushing for full utilization of resources are at most very weak. If, as is plausible, demand-management policy can affect which of the many possible equilibria an economy attains, it can have a lasting effect on the level of output.

Turning to the data, we begin by examining the serial correlation of output over the pre-Depression and post-WWII periods. The time-series properties of output have been a major issue in the rapidly-growing literature on the presence of “unit roots.”⁵ A number of investigators have concluded that output fluctuations are dominated by permanent shocks—that a one percent fall in output

Weir, “Unemployment Volatility: A Sensitivity Analysis,” (Yale University, 1986); and from Nathan Balke and Robert Gordon, “The Estimation of Prewar GNP: Methodology and New Evidence,” *Journal of Political Economy*, (forthcoming 1989).

⁴John Maynard Keynes, *The General Theory of Employment, Interest and Money* (Macmillan, 1936).

⁵For example, John Y. Campbell and N. Gregory Mankiw, “Are Output Fluctuations Transitory?” *Quarterly Journal of Economics*, 102 (November 1987), pp. 857-80.

this year means that a forecaster should revise downward her forecast of output a generation hence by the full one percent or even more. This dominance of permanent shocks has been interpreted as revealing that macroeconomic fluctuations arise not from the demand side but from permanent changes in the economy's production technology.⁶ We show, following John Cochrane,⁷ that substantial persistence in output is largely a post-WWII phenomenon both in the U.S. and in the rest of the world. Before the Depression, "permanent" shocks to output accounted for at most a small part of the year-to-year variance in production.

Some of this increase in persistence might arise from an increased variability of potential and technology growth in the post-WWII period. More plausible in our view is the hypothesis that most of the shift in the serial correlation properties of output arises from successful demand management policies that have largely eliminated the transitory declines in output, caused by movements to inferior equilibria, characteristic of the pre-WWII period.

In order to estimate the magnitude of the improvement in performance relative to potential in the post-WWII period, we construct average output gaps by interpolating potential GNP between major cycle peaks. Even using Romer's data, our output gap measures suggest substantial improvement in performance both in the U.S. and abroad. Weak evidence based on the existence of cyclical asymmetries and the correlations of constructed gaps with observed unemployment suggests that the gaps view may provide a more accurate characterization of fluctuations than the more standard view of fluctuations as cycles around a unique equilibrium trend level of output and unemployment.

The data analysis in the main body of the paper deals with the pre-Depression and post-WWII periods. This omission of the largest transitory fall of output below potential biases the case against finding either a significant improvement in performance in the post-WWII era or evidence that business cycles are best thought of as asymmetric lapses beneath potential, not symmetric fluctuations about trend. In the final section of the paper, we examine the Depression in some detail and reach two primary conclusions. First, the existence of the Depression makes it very difficult to argue that the aver-

⁶Matthew Shapiro and Mark Watson, "Sources of Business Cycle Fluctuations," (NBER, 1988).

⁷John Cochrane, "How Big Is the Unit Root in GNP?" (Chicago, 1987).

age level of output is in any sense the “trend” given by the long-run growth of productive capacity. Either the Depression is a fall below not balanced by any equivalent rise above trend, or the long-term growth rate of the American economy slowed to almost nothing in the decades after WWI. Second, the recovery of output from the Depression carried it almost all the way back to its pre-Depression trend before the stimulus of WWII began to affect the level of U.S. production. Despite greatly reduced capital investment and a substantial labor force withdrawal, the Depression does not seem to have cast a forward shadow on output. This argues for theories of multiple equilibria based on considerations other than the accumulation of human and physical capital.

The paper concludes with discussions of the policy implications of the belief that the economy has multiple equilibria and of the relationship between the multiple-equilibria view and standard approaches to the Phillips curve. Our general conclusion is that the recognition of multiple equilibria suggests the continuing importance of demand management policies, a lesson which is strongly underscored by the continuing European depression.

The proposition that the economy’s equilibrium level of output is inefficiently low is relatively uncontroversial.⁸ To begin with, substantial tax rates on labor income create a wedge between the private and social returns to increased employment. In addition, the pervasiveness of monopoly power in the economy creates some presumption that output is below its efficient level. This presumption is reinforced by evidence that the reservation wages of the unemployed are frequently considerably below

⁸Hall and Taylor, *Macroeconomics*, use a social welfare function that is symmetric in deviations from the natural rate of unemployment. They, however, adopt this only for analytical convenience, arguing (pp. 490-1) that “economists have thought less about the costs of episodes when GNP is above potential. The microeconomic argument supporting the idea that costs are important is the following: The extra work effort needed to push GNP above potential is worth more than is the extra GNP. Instead of working as many hours as they do during a boom and consuming and investing the extra output, the public would be better off with less output and more time to spend on their children, their houses, and in recreation.” But they too believe that “there is at least a range where a boom is socially beneficial even though it is privately costly to workers to be working longer hours.” They attribute this asymmetry to tax distortions that push average labor supply below its optimal level.

going wages, and by arguments involving congestion effects as the unemployed search for work.⁹

It follows that increases in the economy's average level of output and employment would almost certainly be beneficial. The extra output generated would almost certainly more than compensate for any extra wear and tear on capital goods. Demand management policies should, therefore, be used to raise the average level of output—if this is feasible.¹⁰ Whether such an increase is feasible is a major theme in what follows.

CYCLES VS. GAPS

The view that the business cycle consists of repeated transient lapses from potential output is a major piece of the “Keynesian” view: there is often room for improvement, and good policy aims to fill in troughs without shaving off peaks. This Keynesian view stands in opposition to the “natural rate” view: that the business cycle is due to expectational errors that alternately push the economy above and below its sustainable growth path. This natural rate view implies, even in its variants most hospitable to Keynesian concerns, that the scope for macroeconomic policy to affect welfare is small.

To see this, note that the heart of the natural rate view is the claim that the relation between production and prices is well-captured by the stylized Phillips curve relation:

$$(1) \quad \pi_t = \pi_{t-1} + (D_t)$$

where D_t represents the deviation of actual output or employment from a unique equilibrium value determined by tastes and technologies and consistent with steady and anticipated inflation, and where

⁹Arthur M. Okun, “Upward Mobility in a High Pressure Economy,” in Joseph A. Pechman, ed., *Selected Essays of Arthur M. Okun* (M.I.T. Press, 1983), pp. 171-220. See Also Arthur M. Okun, *The Political Economy of Prosperity* (Brookings, 1970); *Prices and Quantities* (Brookings, 1981).

¹⁰Even in new Keynesian models, it may not be. Many menu cost models, for example, have the property that the average level of output is below the optimal level but that all policy can do is iron out fluctuations. See Olivier Blanchard and Nobuhiro Kiyotaki, “Monopolistic Competition and the Effects of Aggregate Demand,” *American Economic Review*, 77 (September 1987), pp. 647-66; also Lawrence Ball and David Romer, “Are Prices Too Sticky?” (NBER, 1987).

inflation in the preceding year is a good proxy for the “anticipated” rate of inflation. Expressions like (1) may be found in leading macroeconomics textbooks.¹¹ The role of lagged inflation in (1) might spring from rational expectations in a context where inflation is nearly a random walk, from adaptive expectations, from long-term nominal wage contracting, or from other forms of nominal inertia.¹²

While references to nominal inertia provide a compelling explanation for unemployment rates that remain above the normal equilibrium level for the length of a business cycle phase, nominal inertia is much less compelling as an explanation for unemployment rates below the natural level. The standard presumption in economic theories with rigid prices is that rationing takes place on the short side of the market: while it makes sense to say that buyers cannot buy or sellers cannot sell as much as they want at quoted prices, it makes much less sense to say that buyers and sellers are forced to transact a greater quantity than they want. The standard Keynesian treatment instead assumes that employment is always demand determined, as if firms could expand employment by compelling workers to accept jobs when unexpected increases in prices reduce real wages.¹³ This illustrates a point we stress below—that fluctuations are more plausibly viewed as adverse shocks pushing the economy below a full-employment equilibrium than as symmetrical fluctuations about a natural level.

The implications of (1) for the efficacy of demand management policy may be seen by summing over time and rearranging:

$$(2) \quad \frac{\sum_{t=1}^T D_t}{T} = \frac{\sum_{t=1}^T 0}{T}$$

Macroeconomic policies which do not raise or lower the inflation rate over a period do not affect the

¹¹ See Stanley Fischer and Rudiger Dornbusch, *Macroeconomics* (McGraw-Hill, 1985), Robert Hall and John Taylor, *Macroeconomics*, and Robert J. Gordon, *Macroeconomics* (Little, Brown, 1987).

¹² See Olivier Blanchard, “Why Does Money Cause Output: A Survey,” (M.I.T., 1987).

¹³ Cf. Hall and Taylor, *Macroeconomics*, p. 490. “Instead of working as many hours as they do during a boom... the public would be better off with less output and more time to spend... in recreation.” But the fact that employment is demand-determined keeps the public from making this choice.

average level of output and employment over that period.¹⁴ Demand management policies mitigate recessions only to the same extent that they choke off expansions. Replacing the lagged inflation rate on the right hand side of (1) with a rational expectation of present inflation would generate a similar conclusion. Expectational mistakes will average out to nearly zero—you can't fool all the people all the time—and the pattern of policy will have no first-order effect on average production. Even were it desirable to increase production on average, such an increase would not be feasible.

Equation (2) leads immediately to a chain of reasoning that concludes that cyclical unemployment should not be the focus of macroeconomic policy. Policies can do no first-order net good or harm on the output side without permanently raising or lowering the inflation rate. We could have avoided the Great Depression—or any other recession that is not followed by an immediate régime shift—only at the cost of incurring a higher steady-state level of inflation today. Why, then, should anyone care about cyclical unemployment? Excess unemployment incurred today because of policy “mistakes” allows a larger boom tomorrow. The business cycle produces welfare losses *only* because consumption is not efficiently smoothed across years.

Robert Lucas has argued that, barring fluctuations of the magnitude of the Great Depression, such welfare losses will be small, might be smaller than the losses from the choice of the wrong long run rate of inflation, and will certainly be orders of magnitude smaller than the losses from policies that retard long run growth—as long as the premise that fluctuations leave the average level of output unaltered is given.¹⁵ The fact that unemployment falls heavily on a few does not justify the Keynesian position. No matter what policy is adopted, the same unemployment total must be divided across years. If long term unemployment is viewed as a special problem, it might even be the case that optimal demand policy requires short, sharp recessions and destabilization—it may be better to leave four people unemployed for six months than one person unemployed for two years—not any form of “leaning against the wind.”

¹⁴In expectational instead of accelerationist formulations, even policies that do permanently raise the inflation rate do not affect the average level of output unless they are unanticipated.

¹⁵Robert Lucas, *Models of Business Cycles* (Basil Blackwell, 1986).

The view that business cycles are fluctuations about supply-determined trends has placed the Keynesian wing of American macroeconomists in a dilemma. They take stylized equation (1) as a starting point, but their acceptance of (1) traps them into fighting for the low ground in their running battles with monetarists: since demand management does not affect average employment, the key issue becomes whether demand affects production for one or three periods in models where the length of a “period” is left unspecified. Concern with the avoidance of excess unemployment as a major aim of public policy can be supported only by a framework that drops the belief that the average is the sustainable level of production.

We are not aware of strong empirical evidence that policy cannot affect the average level of output. Two decades ago, one might have argued that theoretical considerations supported the natural rate hypothesis. It had an arguably superior set of worked-out microfoundations. This is no longer true. A range of theoretical work suggests that business cycles are best approached in terms of lapses from potential output.¹⁶ The “natural” level of output may be like the natural state of a person’s health: desirable, normal—and better than average. One line of work attributes business cycles to collapses of the credit mechanism,¹⁷ and many credit-collapse theories have an affinity with the belief the maximum level of output reached is the optimal level. Recessions arise because contractionary monetary shocks enlarge the effects of failures in financial markets. In such models, it makes no sense to argue that the average trend level of production has any status as *the* sustainable level of output. Banks either remain healthy or they fail. If banks fail there are negative macroeconomic ramifications, but there is no corresponding possibility on the positive side. The appendix presents a highly stylized model illustrating this proposition.¹⁸

Second, models in which aggregate demand externalities allow the economy to settle at any

¹⁶Blanchard, “Why Does Money Affect Output?”

¹⁷See, for example, Bruce Greenwald and Joseph Stiglitz, “Financial Market Imperfections and Business Cycles,” *Brookings Papers on Economic Activity*, 1988 (Spring 1988), pp. ??.

¹⁸Elaborated from N. Gregory Mankiw, “Imperfect Competition and the Keynesian Cross,” *Economics Letters*, 26 (1988), 7-13, and “The Allocation of Credit and Financial Collapse,” *Quarterly Journal of Economics* 101 (August 1986), pp. 455-70.

one of multiple equilibria often have the property that the optimal equilibrium is the one in which the level of production and the rate of resource utilization are highest.¹⁹ Economists have found it relatively easy to construct theories in which aggregate demand externalities are positive and output can settle in inferior low-level equilibria. In addition, many models of hysteresis contain features that keep production from returning quickly to its optimal level once production falls below optimum, but they do not contain features that keep production persistently above the optimal level.²⁰

Third, models in which unemployment arises from sticky prices and rationing fit the belief that output is as a rule below its optimal level. If relative prices are in disequilibrium, it is natural to think that quantities are determined by the desires of those on the short side. Since exchange is voluntary, it is those who wish to exchange the greater quantity who will be rationed. Taking relative price disequilibria as the source of business cycles also leads to the conclusion that output is likely to be often below, but rarely above, its optimal level.

A fourth possible foundation for the potential-gap approach is asymmetric price adjustment. If prices are quick to adjust upwards above previously-anticipated values but slow to adjust downwards, then unanticipated increases in nominal demand will cause inflation but little expansion of production. By contrast, unanticipated decreases in demand will have a larger effect on quantities and a small effect on prices.

By contrast, the microfoundations underlying the trend-cycle approach are limited to theories that view expectational errors as the source of business cycles. A recurring problem facing adherents of the expectational-error view has been the difficulty of pointing to expectational error-generated movements in relative prices large enough to credibly sustain fluctuations of the magnitudes seen in business cycles.²¹ The attempt to evade the difficulty of identifying large real relative price

¹⁹Peter Diamond, "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy* 90 (1987), pp. 881-94.

²⁰For example, Olivier Blanchard and Lawrence Summers, "Hysteresis and the European Unemployment Problem," *NBER Macroeconomics Annual* 1986 (1986), pp. 15-78.

²¹Attempts to shift ground by arguing that firms respond to quantity as well as price signals transform the model into one with

movements by arguing that firms respond to quantity as well as price signals is likely to transform the model into one with imperfect competition and aggregate demand externalities, in which case the average trend level will in general no longer be optimal. In the early 1970's, the expectational-error view could be defended by arguing that its weak points were no larger than would be expected for a productive new research program. But in our view the past fifteen years have seen less progress in repairing these weak points than advocates would have hoped.

These arguments suggest both the normative significance and the actual likelihood of the possibility that demand management policies can affect the average level of output. In order to examine this possibility empirically, we turn next to the comparison of pre- and post-WWII macroeconomic performance.

COMPARING MACROECONOMIC PERFORMANCE

Institutional Differences

There are, of course, tremendous differences between the pre-Depression and post-WWII economies. One set of differences, however, that has attracted notice from prospective and retrospective observers alike²² is the presence in the post-WWII period of that set of “Keynesian” institutions

aggregate demand externalities, in which case the average “trend” level will in general no longer be “optimal.” See Blanchard, “Why Does Money Affect Output?”

²²See Burns, “Progress toward Economic Stability”; and De Long and Summers, “The Changing Variability of Economic Activity in the United States,” for prospective and retrospective views by economists. For a prospective view by one of the most conservative and risk averse securities analysts, see Benjamin Graham, David Dodd, and Sidney Cottle, *Security Analysis: Principles and Technique*, 4th ed. (McGraw Hill, 1962), p. 422: “We believe it reasonable to adopt a somewhat more generous approach to the valuation of common stocks than appeared justified in our previous edition. This conclusion is based on the assurance—not formerly present—of massive Federal intervention to prevent a serious business depression. This now appears to be a basic tenet of both political parties.”

devoted to keeping aggregate demand at a high and stable level. Deposit insurance and a reinforced commitment by the Federal Reserve to its role of lender of last resort acted to stabilize the banking system. The growth of government spending as a share of GNP coupled with progressive taxation gave the government budget a weight as an automatic stabilizer that it had not possessed before. The growth of consumer credit meant that consumer durables purchases were likely to be less adversely affected by income declines. On top of all this came the recognition by governments that they were responsible for preserving high employment and output, and a commitment by governments not to abdicate this responsibility again.

Whether, and which, elements of this institutional complex played a significant role in keeping aggregate demand at a high and stable level during the thirty years since the end of WWII can be disputed. What cannot be disputed is that total nominal demand has been more stable in the post-WWII than it was in the pre-Depression period. The standard deviation of annual nominal GNP growth dropped from 7.8 percent²³ to 3.1 percent; the standard deviation of changes in nominal GNP growth dropped from 9.7 percent²⁴ to 4.3 percent. Impressionistic and statistical evidence both point to a reduction in the demand shocks affecting the U.S. economy. The contrast between the pre-Depression and the post-WWII periods thus provides a good natural experiment for assessing the utility of policies aimed at managing demand.

Data

The data for U.S. output per person of working age used here come from recent reworking of pre-Depression output levels by Romer, deflated by total United States population between 16 and 65 taken from *Historical Statistics of the United States*.²⁵ Nathan Balke and Robert Gordon have also

²³4.8 percent with the WWI period 1915-1921 omitted.

²⁴8.1 percent with the WWI period 1915-1921 omitted.

²⁵Romer, "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1918," and "Gross National Product, 1909-1928: Existing Estimates, New Estimates, and New Interpretations of World War I and Its Aftermath"; Balke and Gordon, "The Estimation of Prewar GNP: Methodology and New Evidence"; *Historical Statistics of the United States*

undertaken to revise the previous estimates of GNP. Throughout the paper we rely on the Romer series because it is the least favorable to our conclusions. It shows less of a reduction in the size of the business cycle across WWII than does the Balke-Gordon series or the standard series. As table 1 shows, both sets of data attribute smaller recessions to the pre-Depression economy than do the original estimates of Kuznets. Figures 1 and 2 plot the Romer, Balke and Gordon, and standard series for GNP per person of working age over the past century.

The Romer series is biased against finding a significant pre-WWII business cycle for yet another reason. It is a series of fitted values adjusted to intercept the standard Kuznets-Kendrick-Gallman series in benchmark years. As a result, it omits any short-run variance in national product not correlated with contemporaneous movements in commodity production. It understates the transitory business cycle component of national product. As Romer acknowledges, in the post-WWII era the components that are omitted from the pre-WWII series account for about a fifth of output variance.²⁶

Two points from figures 1 and 2 are worth noting. The first is the change in the status of the Great Depression as one moves from the Kuznets to the Balke-Gordon to the Romer estimates. In the standard series, the Great Depression is simply the largest of a number of large pre-WWII recessions. The decline in output per person of working age over the post-World War I recession is at least two-thirds, and the declines over the 1893 and 1908 recessions at least one-third, of the magnitude of the Great Depression. According to the standard series, the Great Depression is but the most severe episode of a disease—severe depression—which was endemic in the pre-WWII U.S. By contrast, the Romer series contains no other recession of a magnitude even one-quarter that of the Great Depression.

The second point worth noting is the break in the apparent cyclical pattern before 1890. The decade of the 1880's sees neither large recessions nor labor productivity growth. The change in the apparent character of the business cycle raises the suspicion that the data before 1889 are of much

(GPO, 1975).

²⁶Romer, "The Prewar Business Cycle Reconsidered."

lower quality.²⁷ Accordingly, focus our analysis on the years since 1889.

The unemployment rate data used in subsequent sections consists of the original rate constructed by Lebergott, and the alternatives constructed by Weir²⁸ and Romer.²⁹ Weir's and Romer's estimates are almost surely superior to those of Lebergott, which exhibits excess cyclical volatility.³⁰ The Romer series uses the Lebergott series as raw material and attempts to compensate for excess volatility. Weir tries to build his series up from the raw data while remaining sensitive to assumptions that might introduce excess volatility. All unemployment estimates are highly correlated with each other and with the Romer estimates of output, though the Weir series appears the most highly correlated with the Romer output series.

Sample Periods

We consider two sample periods, one post-WWII and one pre-Depression. The post-WWII period runs over the full range 1947-1987; the pre-WWII period runs from 1889-1929. We exclude the WWII period from consideration because the mechanisms then at work to produce high levels of output reveal little about the peacetime "potential" of the economy. We also exclude the Depression.

The exclusion of the Great Depression biases analysis against finding any significant improvement in economic performance. This may not be appropriate. Perhaps the Depression should be seen as a product of the same economic structure as produced the rest of pre-WWII cycles, in which case the fact that the pre-WWII structure could and did generate such a depression is important evidence of deficiency from the standpoint of macroeconomic performance. Since the Depression is the most virulent outbreak of depression, there may be much to be learned from its study. On the other hand, the Depression is so extraordinary that it dominates statistics over any period in which it is included. Here we exclude the Depression from the analysis; we return to it and consider the lessons

²⁷ See William Shaw, *Commodity Output in the United States* (Columbia University, 1947).

²⁸ Weir, "Unemployment Volatility."

²⁹ Romer, "Spurious Volatility in Historical Unemployment Data."

³⁰ Stanley Lebergott, *Manpower in Economic Growth* (McGraw-Hill, 1964).

of the Depression at the end of the paper.

International Evidence

Historians constructing long run national accounts have quite rightly focused on getting the long run trend right. They have placed first priority on ensuring that the constructed series provide a good picture of the long run sweep of growth. Data that are useful for examining not only long-run growth issues but also the business cycle are rare. In our estimation, there are five industrial economies for which data of reliable quality exist: Canada, France, Germany, Sweden, and the United Kingdom. Long run GNP growth data for these five (and for eleven other industrial nations, including the U.S.) have been constructed and compiled by Angus Maddison.³¹ National product per capita for these five nations is plotted in figures 3 and 4.

The United Kingdom estimates were derived by Feinstein.³² Feinstein's estimates are based on independent expenditure and income series, and are thus probably the most reliable of any of the historical data. Since Feinstein was able to use two independent sources of information on aggregate production. The existence of both income-based and output-based measures of total economic activity his task was considerably easier than that of Kuznets for the U.S. His estimates are correspondingly less likely to suffer from the kind of excess volatility identified by Romer.

Estimates of French production were constructed by Angus Maddison from data in Carré, Dubois, and Malinvaud.³³ The French estimates appear, if anything, to underestimate cyclical volatility. De Long's examination of the path of output during the post-WWI returns to the gold standard suggested that they understated volatility during the 1920's.³⁴

Canadian product estimates are derived from Firestone.³⁵ The Firestone estimates for Canada

³¹ Angus Maddison, *Phases of Capitalist Development* (Cambridge University Press, 1982).

³² C.H. Feinstein, *National Income, Expenditure, and Output in the United Kingdom* (Cambridge University Press, 1972).

³³ Jean-Jacques Carré, Pierre Dubois, and Edmond Malinvaud, *La Croissance Française* (Seuil, 1972).

³⁴ J. Bradford De Long, "Returning to the Gold Standard: Britain and France in the 1920's" (Harvard University, 1987).

³⁵ O.J. Firestone, *Canada's Economic Development: 1867-1953* (Bowes and Bowes, 1958).

may suffer from the excess cyclical volatility diagnosed by Romer in the Kuznets-Kendrick-Gallman estimates for the U.S. The results for Canada should be taken with a grain of salt. Moreover, Canadian macroeconomic fluctuations are closely tied to U.S. fluctuations, and it is questionable that adding Canada introduces much additional real information.

Swedish data are derived from Krantz and Nilsson.³⁶ National product estimates for Sweden at the end of the nineteenth century are thought to be of very good quality for a country as poor as Sweden was then. The precocious development of the Swedish government means that a historian has more government data on production than in most of the other countries. The Swedish data are unlikely to suffer from excess cyclical variability; widespread use of interpolation is more likely to make Swedish data before WWII too smooth.

The German data are the most questionable. They were constructed by Maddison,³⁷ and rely heavily on Hoffman's industrial production estimates. The German national product data are sufficiently less volatile than Hoffman's industrial production data to raise the suspicion that the German data may suffer from deficient volatility.

For all five additional nations the post-WWII period is taken to be 1948-1979. It thus omits the post-1980 period of high unemployment, which we see as involving a shortfall of output from potential of a magnitude unseen since the Great Depression. The comparison is thus between performance before the Depression and performance during the long post-WWII boom. Two prewar periods are used: a pre-WWI period for Britain France, and Germany, all heavily involved in WWI, and the standard pre-Depression period for Canada and Sweden.

THE TIME SERIES PROPERTIES OF OUTPUT

³⁶O. Krantz and C.A. Nilsson, *Swedish National Product 1861-1970: New Aspects on Methods and Measurement* (CWK Gleerup, 1975).

³⁷*Phases of Capitalist Development*, p. 164.

In the past decade, economists have discovered that modern time-series techniques report the existence of a sizeable “unit root” in U.S. output in the post-WWII period. In autoregressive models of output, coefficients on all lags of output sum to one. Fluctuations in production appear not to die away but instead to persist indefinitely.³⁸ One—we believe improper—interpretation of this finding is that business cycle theory has been built on rotten empirical foundations: what is the point of models in which macroeconomic shocks have transitory effects if the canonical output shock never dies away? On this line of thought, the presence of a sizeable “unit root” should lead to the modelling of year-to-year fluctuations as either (i) the result of shifts in “permanent” factors—like tastes and technologies—or (ii) the result of transitory, nominal shocks that nevertheless have permanent “hysteresis” effects.³⁹

Yet on reflection it also appears that finding a strong “unit root” component and failing to find a transitory component in output would be exactly what one would expect if Keynesian automatic stabilizers had substantially reduced the size of the business cycle. The business cycle component of output is naturally thought of as stationary over time, as tending to return to some normal level. By contrast, there is no reason to believe that the potential component of output tends to return to anything that could be called a deterministic trend. Technical progress and capital accumulation do not proceed at constant and deterministic rates. We do not and should not now expect to know what the level of potential U.S. output per person of working age will be in half a century.

Output is therefore the sum of two components, one transitory and one persistent. If the size of the business cycle component shrinks, the persistence of shocks to output will increase, a “unit

³⁸John Y. Campbell and N. Gregory Mankiw, “Are Output Fluctuations Transitory?” and “Permanent and Transitory Components in Macroeconomic Fluctuations,” *American Economic Review Papers and Proceedings*, 77 (May 1987), pp. 111-7; Charles Nelson and Charles Plosser, “Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications,” *Journal of Monetary Economics*, 10 (September 1982), pp. 139-62. But see also Peter Clark, “The Cyclical Component of U.S. Economic Activity,” *Quarterly Journal of Economics*, 102 (November 1987), pp. 797-814; and Mark Watson, “Univariate Detrending Methods with Stochastic Trends,” *Journal of Monetary Economics*, 18 (July 1986), pp. 1-27.

³⁹Blanchard and Summers, “Hysteresis and the European Unemployment Problem.”

root” will become more readily identifiable, and the transitory component may fail to show itself in parsimonious estimated models.⁴⁰ We see the existence of “unit roots” in post-WWII output as a sign that the post-WWII macroeconomic order has performed well and that the business cycle has been small. And our confidence in this interpretation is reinforced by the observation that there is little sign of a “unit root” and much sign of transitory fluctuations in output before WWII—an era which we believe possessed much larger output gaps, and in which output would naturally possess a much larger transitory business cycle component.⁴¹

These points are easily illustrated. Consider a simple model in which output is composed of a (transitory) gap or cycle and a (persistent) potential or trend:

$$(3) \quad y_t^* = y_{t-1}^* + \epsilon_t$$

$$(4) \quad c_t = c_{t-1} + \eta_t$$

$$(5) \quad y_t = y_t^* - c_t$$

where c_t denotes the transitory component. Writing output in terms of its own past values and of shocks ϵ_t that cannot be forecast from output’s past:

$$(6) \quad y_t = (1 + \alpha)y_{t-1} - \alpha y_{t-2} + \epsilon_t - \alpha \epsilon_{t-1} + (1 - \alpha^2)\eta_t$$

where ϵ_t is a serially-uncorrelated shock and $\alpha < 1$ is a parameter that satisfy:

$$(7) \quad (1 + \alpha)^2 = (1 + \alpha)^2 + 2\alpha$$

$$(8) \quad -2 = -2 - 2$$

A fraction:

⁴⁰Policy responses to unanticipated shocks will change the persistence of the transitory component and complicate the story. We have argued that automatic stabilizers have been the most effective Keynesian policy devices, and so focus on the case where demand-management policy leaves the serial correlation properties of the transitory component more or less unaffected. See De Long and Summers, “The Changing Cyclical Variability,” where we also argued that effective automatic stabilizers should lead to a reduction in the persistence of shocks. We now see that argument as misleading; it fails to note that shocks to output in a univariate context include not only business cycle but also long-run productivity components.

⁴¹John Cochrane, “How Big Is the Unit Root in GNP?” (Chicago, 1987).

$$(9) \quad \frac{1 - \rho}{1 - \rho}$$

of the shock to output is permanent. As the relative variance of the business cycle component decreases, the degree to which output is persistent—the “size” of the unit root in output—increases and ρ drops from near one toward zero.

An economist trying to decide whether to model output by (6) or by taking output to be random walk:

$$(10) \quad y_t = y_{t-1} + \epsilon_t$$

might well settle on (10) and conclude that there are no transitory components even when such components contribute the larger share of the variance. For example, suppose $\rho = .75$, $\sigma^2 = .02$, and $\sigma^2 = .01$. For these parameter values, the standard deviation of potential output a decade in the future is a reasonable 3 percent, and the transitory component ϵ_t has a standard deviation of 3.2 percent. Then the error variance from fitting (10) is only five percent greater than the error variance from fitting (6). Only after collecting a sample of 135 years would an economist have a 50-50 chance of rejecting (10) in favor of the alternative (6) at the .05 level.⁴² Yet the transitory component is responsible for 83 percent of the variance in annual output changes. The implication is that a failure of time-series techniques to find transitory fluctuations means not that such do not exist, but only that they do not dominate the data: there are permanent fluctuations mixed in. A change in institutions, or an increase in automatic stabilizers, that reduces the size of the business cycle may therefore lead to the conclusion that there is no transitory cycle.

Moreover, in the context of (3) through (5) estimating business cycle size by examining the variability of output about simple trends may not be fruitful. The stochastic term ϵ_t in potential

⁴²If model selection did not give the random walk (10) the advantage of being the null but were instead based on a more balanced goodness of fit criterion like the Akaike criterion (see H. Akaike, “Information Theory and an Extension of the Maximum Likelihood Principle,” B. Petrov and F Csáki, eds., *Second International Symposium on Information Theory* (Akadémiai Kiadó, 1973)), then the crossover point would come with a shorter sample. The Akaike criterion would give a 50-50 chance of choosing the true model (6) once the sample amounted to 90 years.

growth may contribute a substantial share of the variability of production about simple trends, and so introduce enough noise into variability measures to leave them uninformative about business cycle size.⁴³ For example, suppose that deviations from trend are measured by deviations from a long, exponentially-weighted moving average of output with decay parameter $\lambda = .95$. Then for our sample parameter values, a fifty percent reduction in the size of the transitory business cycle will lead to only a 21 percent reduction in the variability of output about trend.

For an example of a time series for which the transitory components have been severely damped, consider disposable income since WWII. The post-WWII tax and transfer system prevents a considerable share of any transitory fluctuations in GNP from appearing in disposable income. We characterize persistence using simple variance ratios—the ratio of the variance of $y_{t+n}-y_t$ to n times the variance of $y_{t+1}-y_t$. For a random walk all shocks are permanent and all variance ratios equal one. For a pure stationary mean-reverting process the variance ratios will, for high n , approach zero.

For post-WWII disposable income the bias-adjusted two-period variance ratio is 1.08, the five-period .84, and the ten-period 1.01. The ten-period-variance ratio is higher than its expected value if disposable income really were a random walk. Yet the apparent absence of transitory components does not indicate either that all shocks to disposable income are or must be permanent. Instead, it might indicate only that some combination of agencies acts to mask or offset much of what would in other circumstances be a transitory component in disposable income.

The Persistence of U.S. Output Pre-Depression and Post-WWII

If transitory components dominated the sample, the question of their existence would not arise. The inability using post-WWII data to reject the hypothesis that the canonical shock to output is permanent itself suggests that the business cycle component to output has been relatively small since WWII. This inference is reinforced by the dominance of the transitory component for other periods.

⁴³This point becomes more powerful if the random walk of (3) is too simple a model for potential growth. If there is serial correlation in potential growth driven by sustained supply shocks or by decade-long productivity speedup or slowdown, then the risk of contaminating measures of business cycle performance with longer, supply-driven fluctuations becomes larger.

The business cycle of the pre-WWII shows strong signs that fluctuations are as a rule transitory; it is possible to reject the null hypothesis that there was a sizeable unit root in output before WWII.⁴⁴

Table 2 regresses output on lagged output and a time trend and provides strong evidence against the claim that output was close to a random walk before the Depression. Both the null and the alternative are unbelievable in that no one thinks either that output is a random walk or that output is a first-order autoregressive process about a linear trend. Table 2 should be thought of as an attempt to quantify the degree to which output is “close to” a random walk or to transitory fluctuations around a steady trend, not as an attempt to uncover any “true” generating process.

The presence of substantial persistence in output comes through clearly in the post-WWII sample in the large coefficient on lagged output, which is only slightly below its expected value if output per person of working age really were generated by a random walk.⁴⁵ But the pre-Depression sample rejects with ease the null that the coefficient on lagged output was one in favor of the alternative of first-order autoregressive fluctuations about a linear time trend. This conclusion is not sensitive to the yardstick; bias-adjusted variance ratios tell a similar story.

Many have taken the persistence of shocks since WWII to suggest that theories of demand-side fluctuations are fundamentally mistaken. Shapiro and Watson,⁴⁶ for example, write as if whether the univariate process for output is highly persistent will determine the choice between “the Keynesian view, in which fluctuations are predominantly transitory, [and]...the real business cycle view, in which fluctuations are largely the result of permanent shocks.”

The clearly transitory nature of fluctuations before WWII is, in our view, enough to remove

⁴⁴Considering the past century as a whole, it is possible to reject the null that output is a random walk in favor of the alternative of an autoregressive process around a linear deterministic trend. See J. Bradford De Long and Lawrence H. Summers, “On the Interpretation of a ‘Unit Root’ in U.S. GNP” (NBER, 1988).

⁴⁵If the null hypothesis that $y_t = \theta t + 1.0(y_{t-1}) + \epsilon_t$ were true, then the estimated coefficient of y_{t-1} is biased and has the Dickey-Fuller distribution calculated by David Dickey. See Wayne Fuller, *Introduction to Statistical Time Series* (John Wiley, 1976), p. 373.

⁴⁶“Sources of Business Cycle Fluctuations.”

such real business cycle theories from consideration. Shapiro and Watson's Keynesian/real business cycle menu is more likely a choice between concluding that demand management policy was deficient and concluding that demand management policy was more satisfactory. Some set of factors has reduced the relative size of the transitory component in output, and the most plausible candidate is that the business cycle component of production was reduced by the combination of institutions and policies of the post-WWII era. The changing serial correlation pattern of output, in the absence of any strong reason to suspect that the variance of potential growth has become much more unstable, leads to conclude that there has been an improvement in macroeconomic performance—a reduction in the size of transitory components in output—since WWII. In the next section, we assess how much extra output has resulted from this improvement.

MEASURING OUTPUT GAPS

An obvious and simple way to assess macroeconomic performance is to set “potential” equal to actual production at business cycle peaks, and to assume that potential grows at a constant rate between pairs of peaks. The potential series thus generated is piecewise-log linear. Performance can then be assessed by examining the average gap between actual production and this piecewise-log linear peak-to-peak potential series.

Figure 5 plots Romer's series for actual output per person of working age and peak-to-peak potential, where peaks have been chosen by visual inspection before 1930 and by the NBER reference cycle chronology (plus 1966, and omitting 1981) after WWII. The gap for almost any pre-1929 period is much larger than the average post-WWII output gap. The ratio of the average gap over 1889-1929 to that over 1947-1987 is 1.7. Note that the relative stabilization of the post-WWII economy is larger in an gap metric than in a variability metric: the standard deviations of output about quadratic trends in the pre-Depression and post-WWII periods are 3.4 and 2.9 percent, respectively.

Important objections can be raised against the strategy of visually identifying peaks. Peak-to-peak potential suffers from the “weak peak” problem. Not all expansions proceed until the economy’s inventory of unemployed workers is the inventory of the factor of production labor that a benevolent and well-informed central planner would hold. Exogenous shocks interrupt expansions while there is still slack left. Such “weak peaks” should not be given equal weight in the construction of potential.

But which peaks are weak? Using judgment to assess which are the strong peaks creates the possibility that the list of peaks has been unconsciously tailored to fit the desired conclusions. The seriousness of this problem can be gauged from figure 6, which uses a list of “strong” peaks that alternately includes and excludes the 1959 NBER peak. If 1959 is excluded, 1966 ends an extended cycle that is by far the most severe the U.S. has experienced since WWII. But if 1959 is included as a peak, then the average gap is smaller and the growth of potential more variable. Potential grows slowly in the 1950’s, rapidly accelerates in the 1960’s, and decelerates through the 1970’s and 1980’s. The post-WWII period can show better or worse macroeconomic performance depending on the list of peaks chosen.

There are two ways to reduce the impact of the “weak peak” problem. One strategy would use another indicator of slack to check whether a peak is weak. We do not follow this strategy because we wish to use the only effective alternative estimator of slack—the unemployment rate—in later sections to argue for the validity of the potential-gap approach.⁴⁷ The strategy we follow uses the output series itself to assess whether a peak is weak, and leads to a judgment-free constructed potential series.

A Judgment-Free Approach

We use a simple judgment-free approach to construct a family of potential series. Our

⁴⁷Gordon, *Macroeconomics*, uses unemployment rates in benchmark years to calibrate his potential output series to actual production. See also Olivier Blanchard and Danny Quah, “The Dynamic Effects of Aggregate Demand and Supply Disturbances,” (M.I.T., 1987); and Matthew Shapiro and Mark Watson, “Sources of Business Cycle Fluctuations,” (NBER, 1988).

approach is motivated by four assumptions:

- First, potential productivity should not decline over time. Workers and managers do not “forget” production processes.
- Second, actual output is always below potential.
- Third, potential output grows smoothly. New technologies, techniques, and organizations diffuse slowly throughout the economy,⁴⁸ as do changes in natural resources, machines, workers, and tastes for work and leisure. Since the determinants of potential change smoothly, potential itself should grow smoothly as well.
- Fourth, actual output does attain potential on a semi-regular basis. The shocks that reduce output below potential are likely to be damped out in a small number of years.

One could quarrel with these assumptions, but the requirement that output attain potential is close to a definition, and any claim that potential does not grow smoothly because of technology shocks jagged on a year-to-year scale is inconsistent with the picture of “technology” held by economists who have analyzed technical change.⁴⁹ The assumption that potential does not fall is the most vulnerable, for what is an oil or any other negative terms-of-trade shock but a fall in potential? Productive activities that had generated a surplus at previous prices no longer do so. The assumption is that supply shocks like 1973 are rare and their effects small.

The assumption that output is always below potential except during total war is defensible. When employment is high, scarce fixed capital is intensively used, learning-by-doing proceeds at a rapid pace, and unemployment rates for marginal groups in the labor force drop to levels not much worse than the core labor force suffers at business cycle troughs. It is difficult to believe that the inefficiencies of too rapid depreciation or too small “inventories” of unemployed labor are the same order as these benefits of a high-pressure economy.

⁴⁸See David Landes, *The Unbound Prometheus* (Cambridge University Press, 1967); Alfred Chandler, *The Visible Hand* (Harvard University Press, 1978).

⁴⁹Joseph Schumpeter, *Business Cycles* (Harvard University Press, 1939); Simon Kuznets, *Modern Economic Growth* (Yale University Press, 1966); Nathan Rosenberg, *Technology and American Economic Growth*.

The assumption that the economy comes close to potential on a semi-regular basis is more vulnerable. It amounts to assuming that the fluctuations of interest to macroeconomics are those that have the duration of the Burns-Mitchell business cycle.⁵⁰ The persistence of the Great Depression, and of high unemployment in Europe in the 1980's, indicate that this assumption can be unwise. We however, think it justified for studying the U.S. in the pre-Depression and post-WWII periods.

These four assumption inspire the following judgment-free approach for constructing potential output series. Let y^* denote potential, y denote actual output per person of working age, and construct a family of potential series indexed by k using the following recursive procedure:

$$(11) \quad y^*_{t+1} = y^*_t + \max \left[0, \max_{i=1 \text{ to } k} \left\{ \frac{y_{t+i} - y^*_t}{i} \right\} \right]$$

The growth rate of potential between year t and year $t+1$ lies along the straight line with the steepest slope that connects the estimate of potential in year t with actual output per person of working age in any of the years $t+1$ through $t+k$. The assumption that potential growth is smooth implies that the rate of growth from year t to $t+1$ will be close to the rate of growth from t to $t+k$. The assumption that actual is never above potential output requires that potential between t and $t+k$ grows at least rapidly enough to keep potential at or above actual output. And the assumption that actual attains potential on a semi-regular basis implies that for some year between t and $t+k$ actual output is quite close to potential. Different values of the horizon parameter k correspond to different beliefs about the smoothness of potential growth and the frequency with which actual output draws near to potential.

The gap series generated are independent of the average rate of growth of output in the absence of persistent multi-year output declines. But our procedure does not completely eliminate the weak peak problem. Constructed potential output series may well lie below true potential, and estimated average output gaps may be smaller than true average gaps. The average gap measure here should, therefore, be viewed as statistics descriptive of the sample, not as estimates of parameters of underlying generating processes.

⁵⁰Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (Columbia University Press, 1946).

Constructed Potential and Output Gap Series

Generating series for potential output and the output gap for the U.S. confirms the conclusions about the declining size of the U.S. business cycle from the pre-1929 to the post-WWII period reached when potential was generated by simply connecting reference cycle peaks with straight lines. For different values of the horizon parameter k used in constructing potential, the mean output gap constructed is always at least fifty percent greater in the pre-Depression than in the post-WWII period. Figures 7 through 10 present constructed potential based on the Romer estimated GNP series for $k=3$ and $k=5$ in both the pre-Depression and post-WWII periods. Table 4 presents average gaps and standard deviations from quadratic trends for the Romer series. The mean gap is about 50 percent larger in the pre-Depression than in the post-WWII period; the amount of macroeconomic slack proportional to the size of the economy was some 50 percent greater before the advent of Keynesian demand management policies and the post-WWII framework of international trade.

This difference in mean output gaps suggests that the United States has on average come one percent of production closer to potential output since WWII than it did before the Depression—a surplus that, if sustained, would now run at almost \$50 billion a year. Since this improvement in performance is to a first approximation a simple increase in capacity utilization, it is almost all a pure welfare gain. If there is an association between a high-pressure economy on the one hand and rapid growth of potential productivity on the other, then the benefits from improved macroeconomic performance may be substantially understated.

The superiority in estimated performance of the post-WWII period in the United States does not come equally from all post-WWII decades. The improvement stems largely from reasonably good performance in the 1950's and excellent performance in the 1960's. Pre-Depression average gaps are not fifty but a hundred percent larger than gaps during the first half of the post-WWII period, and the difference between pre-Depression and post-1970 average gaps is only 30 percent.

Such a difference in the size of the average gap is striking given the near-equality of the volatility-based estimates of business cycle size, the standard deviations of levels and growth rates from quadratic trends, also reported in table 4. In light of the changing persistence properties of out-

put and the large reduction in the mean constructed output gap, we read the near-equality of the standard deviations of output from trend as suggesting that volatility measures have been corrupted by noise introduced by the stochastic nature of long-run potential output growth. The substantial reduction in the volatility of year-to-year changes in output reinforces our belief that the small size of the reduction in volatility about trend from the pre-Depression to the post-WWII period springs from noise introduced into this variability measure by stochastic potential growth.

The fact that measures of volatility about trend stand alone in opposition to many other pieces of evidence—the changing persistence properties of output, the reduction in the volatility of year-to-year changes, and the substantial reduction in business cycle size when calculated in a gap metric—casts doubt on the validity of conclusions that spring from comparisons of U.S. output volatility about trends. Since the revisionist interpretation of relative macroeconomic performance carried out by Romer depends in substantial part from the use of volatility as a proxy for performance, we see no reason to depart from our belief that the post-WWII period has seen a large improvement in macroeconomic performance.

The difference between the conclusions about the relative sizes of the pre-Depression and post-WWII business cycle that emerge from Romer's comparisons of volatilities and our comparisons of gaps demonstrate that whether one believes that business cycles are lapses from potential output or cycles around trend output makes a difference not only for whether one thinks that policy can significantly improve macroeconomic performance but also for whether one thinks that different eras have seen significantly different levels of macroeconomic performance. Cycle-based measures produce no warrant for believing that the Romer series shows that macroeconomic performance has been significantly better since WWII than it was before the Depression. By contrast, gap-based measures strongly support the claim that the post-WWII U.S. economy has—for Keynesian or other reasons—seen superior performance. The gap-based measures mesh with the conclusions implied by the serial correlation properties of output. This reinforces what we see a strong *a priori* case for gap-based measures.

Moreover, there is weak empirical evidence that the gap approach is superior to the cycle approach. Okun's law appears stronger in a gap- than in a cycle-based framework, and the distribution of constructed output gaps appears skew in a way that would be expected if the generating process were of the "gaps" type but is hard to rationalize if the generating process were of the "cycles" type. These considerations, however, are weak. They may give some added confidence to a reader already sympathetic to this approach; it will not convince skeptics.

Fluctuations in the unemployment rate are an alternative gauge of business cycle fluctuations. To the extent that gaps or cycles really are measures of business cycle components of output, they should be correlated with fluctuations in the unemployment rate. A higher correlation of gap than cycle measures with unemployment would suggest that the constructed output gaps were a better measure of the business cycle than the constructed cycle measures. Indeed, a number of possible unemployment rates appear to be more highly correlated with the constructed output gap than with measures of cycles around trend.

For pre-WWII unemployment rates, we use the original series of Lebergott, the replacement constructed by Romer, and the alternative replacement series constructed by David Weir, who attempts by returning to original sources to overcome the problem of excess volatility originally pointed out by Romer. We first compare how the constructed gap ($k=5$) as a predictor of unemployment fares vis-a-vis output itself and a time trend. If the proper measure of the state of the business cycle is indeed the deviation of output from a near-deterministic trend, the output gap measure should add no additional explanatory power. Yet the output gap is a better predictor of unemployment.

This result might arise simply from the fact that our potential series capture a relatively flexible trend. Perhaps "cycle" measures, if allowed a similarly flexible trend, would also perform as well as predictors of unemployment. While we would not be surprised to learn that there is a method of estimating an average trend which does generate a business cycle series as highly correlated with unemployment as the constructed gap, to date we have not found such a method. We have used as cycle measures the deviation of output from long moving averages of output, from exponentially smoothed output (with smoothing coefficient equal to .7, which leads to a better fit than any other trend series tried to date), and from a trend series constructed using the non-linear filter proposed by Hodrick and

Prescott.⁵¹ The output gap appears a better predictor of the unemployment rate than the cycle-based measures for all three of these proposed flexible average trends—although the difference is statistically significant at the .05 level only for the Weir unemployment series.

In the cycle-trend approach, all years carry approximately equal amounts of information as to what the sustainable level of production is. In the potential-gap approach, only peak years carry much information about the sustainable level of production. The fact that gap-based outperform cycle-based measures in predicting unemployment suggests that peak years carry much more information about what the sustainable level of output is than do average years. Business cycles are, it appears, better thought of as falls away from peaks than as symmetric fluctuations about trend.

These regressions, however, provide only weak evidence. While the difference between the coefficients on the gap and cycle variables is always of the proper sign and at least as large as one standard deviation, the difference remains statistically significant for all ways of constructing the cycle only for the Weir unemployment rate series. If one found the Weir series less trustworthy than the Romer series, the natural conclusion would be that we really cannot say whether gap measures are truly superior to cycle measures as predictors of unemployment.

Skewness

If business cycles are generated by expectational errors, there is some reason to expect the distribution of production to be symmetric. The production of a downwardly-skewed distribution of output would require a downwardly-skewed distribution of expectational errors. This seems unlikely: there is no immediately apparent reason for underestimates of money growth to be small and frequent and overestimates to be large and infrequent, as would be required to generate skewness in the distribution of expectational errors.

⁵¹Robert Hodrick and Edward Prescott, "Post-War U.S. Business Cycles: An Empirical Investigation" (Carnegie-Mellon University, 1980). The filtered series is affected by the treatment of the endpoints 1889, 1929, 1950, and 1987. We anchor the filtered trend to actual output at the endpoints. The smoothed series thus does not "see" the approaching Depression. If trend did "see" the Depression in advance, the cycle would be a horrible predictor of unemployment.

By contrast, in a “gaps” framework, skewness of output is natural and expected. Production cannot go above, but it can—in a bad enough recession—drop arbitrarily far below potential. The combination of many small output gaps and a few large ones will give a skew to the distribution of lapses from potential output. As table 8 shows, for small horizons k the empirical distribution of calculated output gaps is more skew than would be likely if the generating processes were ARIMA’s with normal innovations.⁵²

Because the procedures we use to construct our gap measures force potential to be high, they will introduce some skewness into output even when the underlying distribution is symmetrical. Nevertheless, significance levels calculated from Monte-Carlo simulations indicate that for the past century (excluding the Great Depression and WWII) the actual distribution of constructed output gaps is abnormally skew. Including the Great Depression would greatly strengthen the rejection of the null that the actual distribution of output gaps is no more skew than would be expected from a symmetric generating process. These skewness-based tests, however, provide at best a weak rejection of symmetric “cycles” models. For neither the pre-Depression nor the post-WWII period taken individually is the skewness of output gaps statistically significant, and even for the entire sample output gaps are abnormally skew only if they are constructed using small values of the horizon parameter.

Output is not the aggregate variable that shows the most signs of business-cycle asymmetry. In the 1982 recession the U.S. unemployment rate rose 4-5 points above its average level; it is unthinkable that it would fall 4-5 points below its average level. The unemployment rate has been found in a number of papers to exhibit a significant degree of skewness toward positive rates.⁵³

⁵²Asymmetry of deviations of output from a stochastic trend has previously been found by Daniel Sichel, *Four Essays on Business Cycle Asymmetry* (Princeton University Ph. D. Diss., 1988). In an earlier paper, we tried and failed to find evidence of asymmetry in output growth rates or in deviations from simple polynomial trends. We now attribute our failure to find asymmetry to a failure to realize the implications of stochastic potential growth for the tests we previously attempted. See J. Bradford De Long and Lawrence H. Summers, “Are Business Cycles Symmetrical?” in Robert J. Gordon, ed., *The American Business Cycle: Continuity and Change* (Chicago, 1986), pp. 166-78.

⁵³De Long and Summers, “Are Business Cycles Symmetrical?” and Salah Neftçi, “Are Economic Time Series Asymmetric Over

There are a few years in which the unemployment rate is much higher than its average value, and many years in which the unemployment rate is a little bit lower than its average value. For the post-WWII period with its small transitory components to fluctuations, the unemployment rate exhibits a minimal but positive skewness of .23. And for the pre-Depression period the unemployment rates are highly skewed indeed, with skewnesses of .56 for the Weir series, 1.34 for the Romer series, and 1.32 for the Lebergott series. While it is possible that this asymmetry in unemployment springs from nonlinear adjustment dynamics in the labor market, we believe unemployment asymmetry is more likely to be a confirmation of the somewhat weaker asymmetry we detect in output relative to potential.

The evidence on business cycle asymmetry, however, remains relatively weak because measures of asymmetry are by definition sensitive to outliers. Confidence intervals vary widely depending on the exact shape of the underlying distribution. Evidence on asymmetry, like evidence on the superior fit to Okun's law provided by gap measures, provides only weak empirical support for the gap point of view.

International Evidence

Examining the long-run business cycle performance of other economies by and large confirms the conclusions we have reached from our examination of the United States. Output becomes more persistent after WWII, suggesting a reduction in the size of the transitory component of output. In addition, output gap measures of business cycle size also show a considerable drop across the pre-Depression/post-WWII interval. The association of decreased output gaps with increased persistence to output also holds for all countries except one. Calculated variance ratios show that in all nations except Britain output fluctuations have been more persistent since WWII than they were before. The exception, Britain, exhibited remarkably steady output growth from 1948 to 1979, with few if any permanent shocks to the level of per capita output.

An examination of other nations thus confirms the broad picture gained by study of the U.S.

In four of five nations post-WWII period has indeed seen an increase in the persistence of output fluctuations, suggesting a decline in the relative magnitude of the transitory “business cycle” component of output. Similarly, in four of five nations post-WWII output gaps are noticeably reduced.

THE GREAT DEPRESSION

The preceding sections have dealt with the pre-Depression and post-WWII periods only. They have left the Great Depression to one side. In so doing, they have not considered the episode that argues most powerfully both that output fluctuations are transitory and that the average level of output cannot be identified with any supply-driven “trend.” The Depression is hard to reconcile with the view that the average level of output over time reveals what the sustainable level of output is, for it is next to impossible to make sense of U.S. long-run growth without assuming that the 1929 peak carries a lot more information about productive capacity than does the 1933 trough. It is impossible to imagine an episode that would push output as much above as the Depression pushed output below its average level. The Depression, by its very existence, makes it impossible to argue that fluctuations around averages are symmetrical.

The Depression also provides a reason for believing in multiple equilibria as opposed to some hysteresis alternatives to the natural rate view. The Depression was transitory: at its end, the U.S. economy was all the way back to the pre-Depression trendline projected forward from 1929. And if the channels of influence involving reduced human and physical capital accumulation had little effect then, it is difficult to think of circumstances in which they would have significant effects.

Mean Reversion and the Great Depression

Examination of the Great Depression reinforces the belief that output shocks before WWII were transitory. The Depression reinforces the belief that the economy may possess multiple levels of production that are equilibria in the sense that the forces pushing output back to potential are weak. The Depression provides little support for the position that “the time path to equilibrium shapes...

equilibrium.”⁵⁴

The output decline of the Great Depression was not permanent. Output did recover. The log of per person of working age output declined by .383—.455 relative to the 1889-1929 trend—between 1929 and 1933. But between 1933 and 1941 the log of output per person of working age grew by .529—.384 more than the trend. By the time WWII began and the government began to exert significant degrees of direct command over the economy, more than five-sixths of the Depression decline in output relative to trend had been made up. It is difficult to attribute any of the pre-1942 catch up of the economy to the war. As figure 11 shows, neither the federal government’s fiscal deficit nor the surplus on trade account became an appreciable share of national product before Pearl Harbor.

Figure 12 below plots the path of output per person of working age before and during the Great Depression. For reference three trend lines are given. The first, 1906-1929 trend line is the upper linear envelope of the points attained before the Great Depression. The second, 1892-1929 trend line is meant to capture the long-run average growth rate of per person of working age output. And the third, 1919-1929 trend line is meant to capture any post-WWI acceleration in productivity growth. All three of these trend lines fall close together. Taking one’s bearings from any of them, the degree to which output had recovered from its Depression depths even before the United States entered WWII was substantial.

The substantial degree of mean reversion exhibited by 1941 is a strong piece of evidence that shocks to output are transitory. Suppose that U.S. production really were a random walk: a one percent surprise decline in output today should lead to a one percent downward revision of one’s estimate of production in the far future. Then the decline and subsequent rapid recovery of the 1929-1941 period are orders of magnitude more unlikely than even a decline of the magnitude of the Depression itself.

Calculating the probability under the null hypothesis that output follows a random walk of a Great Depression—a 45.5% decline in output followed by a 38.4% recovery relative to previous drift

⁵⁴Edmund S. Phelps, *Inflation Policy and Unemployment Theory* (Macmillan, 1972), pp. 77-8.

within a twelve-year period—is straightforward.⁵⁵ But such a calculation alone is not sufficient. The Depression itself draws our attention to the decade 1929-1941 in the first place. The appropriate way to determine the probability of a reversion to mean as complete as exhibited by the Depression is to calculate the likelihood of a Depression—of such a decline and recovery—provided that events over 1929-1941 were such as to draw our attention to this decade. Suppose, conservatively, that *only* a decline in output relative to trend as large or larger than actually experienced in the Depression would draw our attention to the decade. Divide the probability of a Depression-sized episode that contains such a substantial degree of rapid mean reversion by the probability that our attention will be attracted to the 1929-1941 period by a decline at least as large as the Depression. This probability, reported in table 11, reveals that mean reversion in the Depression alone is sufficient to justify rejection of a random walk null even after correcting for the fact that our decision to examine the Depression is based on the fact that it is an extraordinary episode. Table 11 assesses the strength with which the Depression can reject a random walk null for three different estimates of the standard deviation of year-to-year changes in output.

The recovery of output to within 7.1% of pre-1929 trend by 1941 is unlikely if output were a random walk, even after correcting for the fact that our choice of 1929-41 as the period to examine is motivated by the fact that it was extraordinary. The recovery of output from 1933 to 1941 is by itself more than enough evidence to reject the null hypothesis that output shocks are in general permanent at conventional significance levels.⁵⁶

Full recovery from the Great Depression weighs strongly in favor of multiple-equilibrium as opposed to hysteresis alternatives to the natural rate hypothesis. Models in which the economy can

⁵⁵John Harrison, *Brownian Motion and Stochastic Flow Systems* (Wiley, 1985).

⁵⁶The decade 1929-1939 is not a draw from a mean-reverting stochastic process that characterizes the rest of the pre-WWII period. If the same autocorrelations did prevail, the Depression would have been over by 1935: no other business cycle lasted for more than six years. But this is an argument for the line of research followed by Stock that allows each business cycle to evolve on its own time scale, not for any conclusion that declines in output are in general permanent. See James Stock, "Measuring Business Cycle Time," *Journal of Political Economy*, 95 (December 1987), 1240-61.

fall into a low-production, low-activity state and remain there for considerable periods of time with no noticeable tendency to return to equilibrium are not inconsistent with the Depression. Models in which falls in output and increases in unemployment themselves have strong effects on the natural rate of unemployment and natural level of output are not. The Depression as a whole does not seem to have materially reduced the long run growth path of the U.S. economy.

Full recovery from the Depression is particularly striking given that so many of the mechanisms that economists rely on to produce “hysteresis” in models were at work during the Depression. Net capital formation was nil during the Depression: reproducible tangible assets in 1939 had the same real value as in 1929.⁵⁷ Labor force growth slowed: one aim of the Social Security program as instituted was to provide incentives for workers to retire and so reduce excess supply in the labor market. In this aim the program was successful. Labor force participation among men 65 or older, which had drifted downward from 68 to 55 percent from 1890 to 1930, had dropped to 41 percent by the 1940 census and continued to fall thereafter.⁵⁸ The failure of hysteresis effects to emerge on the aggregate supply side as a result of the Depression when the channels through which such effects are presumed to work were active suggests that it is appropriate to view the economy in general not as a system in which the path to equilibrium affects the position of equilibrium but as a system with multiple short run equilibrium positions.

CONCLUSIONS

To sum up, we list four principal conclusions. First, the failure of time series analysis to find strong evidence of a transitory business cycle in the post-WWII U.S. economy is an argument that recent macroeconomic performance has been good. Before WWII, no one doubts that production contained transitory “business cycle” components. The Great Depression alone provides sufficient evidence to reject the claim that the canonical shock to production is a permanent one. Changes in

⁵⁷ *Historical Statistics*, p. 256.

⁵⁸ *Historical Statistics*, p. 132.

potential output are presumably permanent and persistent, while changes in the output gap are likely to be ephemeral and transitory. A reduction in output gaps will change the serial correlation properties of output and may leave time-series analysis unable to identify a transitory component. The lack of a transitory component in output since WWII—the finding that the canonical shock to production is persistent in a univariate context—tells U.S. that performance has been good and that the shocks that used to produce the business cycle have been damped by a robust economic structure or by skillfully-conducted policy.

Second, both theoretical and empirical considerations weigh on the side of gap- as opposed to cycle-based decompositions of output. Okun's law is more of a law if output benchmarks are gap-based measures of potential instead of cycle-based measures of average trend. Gaps are good indicators of performance even if the economy is well-described by a trend-cycle model, but cycles may not be good indicators of performance if the proper description of the economy focuses on the gap. Moreover, present-day microfoundations are more in sympathy with approaches based on gaps than on cycles.

Third, assessing performance by gaps establishes a sizeable relative improvement in performance since WWII. The significance of this improvement had been cast into doubt by Christina Romer's improved estimates of macroeconomic aggregates, which show only a small decline in volatility between pre-Depression and post-WWII periods. Although Romer's work does greatly reduce the apparent size of pre-Depression business cycles, the most extreme interpretation of Romer's work—that there has been no significant decline in the business cycle between pre-Depression and post-WWII periods—does not survive the removal of the assumption that volatility is an adequate measure of performance.

Fourth, the most important feature of the Great Depression is that the U.S. recovered from it, that the Great Depression appears to have had few effects on the long-run growth path of the U.S. economy. We read this as a sign that the metaphor of "hysteresis" as applied to economies should be understood as asserting not that there are no tendencies after demand shocks for the self-regulating mechanisms of the economy to push unemployment down and output back to trend levels, but that there is likely to be a wide range over which the self-regulating mechanisms prove to be weak.

These conclusions suggest the likelihood that successful efforts to manage demand might well have increased average output and reduced average unemployment during the post-WWII period. It is very unlikely that the sharp contractions of the pre-WWII period had substantial inflation dividends: we see no evidence that booms reached higher peaks before WWII because the ever-present danger of deep depression made the absence of such contractions a large expansionary surprise.

It is less clear whether the same is true of the post-WWII period. In Europe, the highly contractionary demand management policies of the early 1980's have left a legacy of higher unemployment and lower output, which have to date already lasted an order of magnitude longer than any believer in the natural rate hypothesis would have predicted a decade ago. A plausible case can be made that many of the structural problems Europe faces today are in fact the consequences of the dramatic increases in unemployment that occurred during the disinflation of the 1980's. Had the disinflation proceeded more gradually, or had inflation been permitted to remain steady, output and employment might well be much higher today. The contraction of the early 1980's appears to have knocked European output down to an unfavorable real equilibrium, and in the absence of rapid demand expansion there appear to be no strong push back toward full employment, just as there were no strong forces pushing Britain to full employment in the 1920's or the U.S. to full employment in the 1930's.

The real issue for the U.S. today is whether it would have been possible to have increased output in years like 1958, 1975, and 1982 without having given rise to a legacy of permanently higher inflation. This possibility is obviously ruled out by mechanistic Phillips curves. A belief in multiple real equilibria, however, leaves no space for such a Phillips curve. We note that the U.S. has suffered only one severe recession of the magnitude relatively common in the pre-WWII era—1982—in the past forty years. We do not see that this avoidance of what in the pre-WWII era would have been seen as the inevitable share of severe recessions has carried with it any reduction in the levels of output reached in booms. And inflation now is essentially identical to what it was immediately before the Korean War. We believe that the answer to whether the U.S. could have avoided 1958, 1975, and 1982 without ever-accelerating inflation is that the U.S. has successfully avoided all but one of the post-WWII equivalents of the 1894, 1908, and 1921 troughs, and yet has not avoided the post-

WWII equivalents of the 1901, 1906, and 1929 peaks. Whether still further significant improvement would have been possible is a much more uncertain prospect.

APPENDIX: A SIMPLE MODEL OF GAPS

Following Mankiw,⁵⁹ consider an economy with one representative consumer, with utility given by:

$$(A1) \quad U = C^{\alpha} S^{1-\alpha} (1-\ell)^{1-\beta}$$

where ℓ is the fraction of time spent at work, C is real consumption, and S is saving. Leisure ($1-\ell$) and consumption (C) are goods in themselves; since we see business cycle-related unemployment as almost entirely social waste, we consider $1-\beta$ close to zero. Saving (S) is valued because it promises to increase command over resources in the future.

The consumer allocates time between work and leisure, and allocates income between expenditure on consumption goods (pC , where p is the price of material output) and expenditure on investment (S , saving). She also receives profit income π from her ownership of the economy's firms. Her budget constraint is given by:

$$(A2) \quad pC + S = w\ell + \pi$$

where the wage is numeraire and ℓ is both the amount of work performed and the consumer's labor income. Maximizing utility leads to the consumption function:

$$(A3) \quad C = \frac{(1+\beta)w\ell + \pi}{p}$$

and to the desired level of savings:

$$(A4) \quad S = (1+\beta)w\ell + \pi - pC$$

The production side of the economy is composed of a large number N of firms each selling a slightly different product. We neglect entry and set demand for each of the N products at:

$$(A5) \quad D_i = \frac{Q}{N} \left(\frac{p_i}{p} \right)^{-\frac{\mu+1}{\mu}}$$

where p_i is the price of the i th product. Each firm produces using a constant returns to scale technology with labor as the only input:

⁵⁹Mankiw, "Keynesian Cross."

$$(A6) \quad Q_i = \dots$$

firms' total production and employment sums to total material output and employment in the economy:

$$(A7) \quad Q = \sum_{i=1}^N Q_i$$

In equilibrium each firm produces:

$$(A8) \quad Q_i = \frac{Q}{N}$$

and charges a price:

$$(A9) \quad p_i = 1 + \mu$$

The general price level is then also:

$$(A10) \quad p = 1 + \mu$$

And profits, the surplus over variable labor cost received from the sale of goods, are:

$$(A11) \quad \dots = \mu Q$$

The model diverges significantly from Mankiw's "Keynesian Cross" and follows closely the line of his "Allocation of Credit and financial Collapse" in the financial sector. The investment-financial sector is plagued by adverse selection. Ultimate savers commit their savings to banks, which then commit them to entrepreneurs to use in investment projects. There are S/p entrepreneurs, each of whom has unique knowledge of a different investment project i and its probability of success π_i . Probabilities of success are distributed uniformly over $[0, 1]$. Each project has the same expected gross return $2 > R^* > 1$, and each requires an irreversible purchase of 1 unit of the manufactured good.

Banks can commit their principals' funds to several different opportunities. They can loan money to entrepreneurs making risky but productive investments. They can loan money at the risk-free rate to other banks that see better investment opportunities. Or they can hire labor to engage in rent-seeking activities—employing lawyers to try to take away others' property rights, or hiring brokers to trade on inside information that will soon be publicly announced. Rent-seeking activities yield a private, but not a social, marginal product of unity. Competition will then drive the risk-free inter-bank loan rate to one also.

The loan market is segmented into many little loan markets, each one competitive on the banking side. Entrepreneurs in market j borrow from private banks and commit themselves to repayment at the bank's set interest rate of R_j if their project is successful. An entrepreneur in loan market j will borrow from the bank if her personal r_i is such that:

$$(A12) \quad \frac{R^*}{r_i} > R_j$$

The average probability of repayment in market j is a decreasing function of the market interest rate R_j :

$$(A13) \quad \frac{\text{avg}}{j} = \frac{R^*}{2R_j}$$

The small loan markets are of two types. The first consists of markets in which entrepreneurs are fortunate enough to be liquid. They hold government-guaranteed paper or otherwise insured liquid assets of quantity S that they can use as collateral for their borrowings. Consider the case in which $S > 0$. In this case loans to entrepreneurs in this first class of markets cease to be risky, and banking competition drives the equilibrium rate R_j to one, the private value of funds in the alternative, rent-seeking activity. The amount of real capital investment financed through this segment of the loan market will be S/p .

Now consider the case in which $S = 0$. Banks in this first sector will borrow on the interbank market and commit all society's savings to investments, and the risky interest rate will be such to clear the market. The amount of real capital investment financed through this segment of the loan market will be S/p .

The second type consists of markets in which entrepreneurs not have access to insured or guaranteed credit to serve as collateral. For a given average probability of successful investment in these markets, competition among banks will drive the local interest rate to:

$$(A14) \quad R_j = \frac{1}{\frac{\text{avg}}{j}}$$

This second class of loan markets has no equilibrium. Since $R^* < 2$, the probability of repayment at any given market interest rate is too low for banks to make a profit at that interest rate. No loans are

made, and all desired savings committed to banks in this second, unfortunate class of local financial markets are either loaned at the safe interbank rate of 1 to banks in the first class, or spent on churning portfolios and other rent-seeking activities. The actual quantity I of output purchased for investment and installed as new capital is therefore:

$$(A15) \quad I = \frac{\min(S, \quad)}{p}$$

Goods market equilibrium—that consumption C plus investment I sum to total material production Q —closes the model. If (A15) does not bind and the investment-financial sector has sufficient liquidity to support all desired investments, then material production equals:

$$(A16) \quad Q = \frac{+}{1 + \mu(1 - (+))} \quad \text{if } \frac{-}{p}$$

If (A15) does bind, then material production depends on the liquidity of the economy:

$$(A17) \quad Q = \frac{+}{1 + \mu(1 -)} \quad \text{if } \frac{-}{p}$$

These levels of material output are lower than the level of output that a social planner would choose:

$$(A18) \quad Q^* = +$$

In the case where there are no benefits from increased leisure—where reducing working time below the standard working day is a total social waste and $1 - - = 0$ —then (A16) through (A18) become:

$$(A16') \quad Q = 1 \quad \text{if } \frac{-}{p} 1-$$

$$(A17') \quad Q = \frac{+}{1 + \mu(1 -)} \quad \text{if } \frac{-}{p} 1-$$

$$(A18') \quad Q^* = 1$$

In this case, the presence of market power in product markets does not by itself push the economy below capacity as long as there are no financial market failures. The existence of market failures in the investment-financial sector is the key to reducing material output below capacity. And material output is never above capacity, understood as the level of output that a planner would choose.

This simple model has glaring defects. It is static and it assumes that the labor market always clears, to name two. Nevertheless, it exhibits a multiplier: a unit increase in the real liquidity of the economy increases material output by a multiplier that approaches $1 + \mu$ as the marginal propensity to consume approaches one. And it illustrates that business cycles driven by liquidity or coordination failures can lead to lapses below, not fluctuations around, sustainable levels of output.

TABLE 1
FALLS IN GNP PER WORKING AGE ADULT DURING RECESSIONS, 1889-1929

Recession	Romer	Output Series:	
		Balke-Gordon	Standard
1892-4	-.080	-.079	-.126
1895-6	+.002	-.045	-.047
1901-2	-.006	-.007	-.015
1906-8	-.073	-.118	-.117
1913-5	-.053	-.098	-.097
1918-21	-.084	-.146	-.231
1926-8	-.013	-.011	-.029

Source: Romer series, Balke-Gordon from "Estimation of Prewar GNP," and standard series from *Historical Statistics*.

TABLE 2
PRE-DEPRESSION AND POST-WWII OUTPUT PERSISTENCE

Variable	Period	Equation	Residual	
			Std. Dev.	1st Autocorr.
Output per Person of Working Age	1889-1929	$Y_t = .010(t) + .373(Y_{t-1})$ (.002) (.152)	=.031	$t_{t-1}=.11$
Output per Person of Working Age	1947-1987	$Y_t = .003(t) + .816(Y_{t-1})$ (.002) (.081)	=.024	$t_{t-1}=.08$

Source: Authors' calculations using the Romer series.

TABLE 3
BIAS-ADJUSTED VARIANCE RATIOS

Variable	Period	Horizon:		
		2	5	10
Output per Person of Working Age	1889-1929	.911	.429	.194
Output per Person of Working Age	1948-1987	1.151	1.018	.866

Source: Authors' calculations using the Romer series.

TABLE 4
PEAK-TO-PEAK OUTPUT GAPS FOR VARIOUS HORIZONS: ROMER SERIES

Period	3 Yr	Horizon		Std. Dev. of Output from Trend of Changes	
		5 Yrs	8 Yrs		
Post-WWII	.016	.023	.028	.029	.026

Pre-Depression .026 .035 .043 .034 .038

Source: Authors' calculations using the Romer series.

TABLE 5
REGRESSIONS OF UNEMPLOYMENT ON THE OUTPUT GAP AND OUTPUT
1889-1929 and 1948-1987

Unemployment Series	Coefficient of			Rho	SEE
	Output Gap	Output	Time		
Weir	-.565 (.053)	0.040 (.040)	-.053 (.056)	0.69 (.08)	0.90
Lebergott	-.838 (.092)	0.167 (.055)	-.289 (.096)	0.78 (.07)	1.57
Romer	-.541 (.062)	0.069 (.037)	-.117 (.065)	0.80 (.07)	1.06

Source: Authors' calculations using unemployment data from Romer's "Spurious Volatility," Weir's "Unemployment Volatility," and Lebergott's *Manpower in Economic Growth*, and using the Romer series for output.

TABLE 6
UNEMPLOYMENT ON OUTPUT GAPS AND MOVING-AVERAGE CYCLES 1889-
1929 and 1948-1987

Unemployment Series	Coefficient of:			Rho	SEE
	Output Gap	9-Yr. Moving Avg. Cycle	Time		
Weir	-.499 (.144)	-.042 (.128)	0.011 (.004)	0.63 (.09)	0.94
Lebergott	-.515 (.244)	-.188 (.218)	-.009 (.008)	0.76 (.08)	1.59
Romer	-.391 (.166)	-.100 (.148)	-.004 (.005)	0.79 (.07)	1.09

Source: Authors' calculations using unemployment data from Romer's "Spurious Volatility," Weir's "Unemployment Volatility," and Lebergott's *Manpower in Economic Growth*, and using the Romer series for output.

TABLE 7
UNEMPLOYMENT ON OUTPUT GAPS AND EXPONENTIALLY-SMOOTHED
CYCLES 1889-1929 and 1948-1987

Unemployment Series	Coefficient of:			Rho	SEE
	Output Gap	Exp. Smoothed Cycle ($\alpha=.7$)	Time		
Weir	-.548 (.131)	0.031 (.179)	0.016 (.004)	0.68 (.08)	.92
Lebergott	-.361	-.402	-.007	0.78	1.59

	(.225)	(.305)	(.007)	(.08)	
Romer	-.456 (.152)	-.015 (.206)	0.002 (.005)	0.80 (.07)	1.08

Source: Authors' calculations using unemployment data from Romer's "Spurious Volatility," Weir's "Unemployment Volatility," and Lebergott's *Manpower in Economic Growth*, and using the Romer series for output.

TABLE 8
SKEWNESS OF CALCULATED OUTPUT GAPS

Period	Horizon	Skewness	Marginal Significance Level in:	
			Levels Model	Differences Model
Pre-Depression & Post-WWII	k=3	1.59	.03	.04
	k=4	1.10	.12	.15
	k=5	0.77	.30	.40
Pre-Depression	k=3	1.36	.07	.09
	k=4	0.99	.20	.20
	k=5	0.66	.50	.50
Post-WWII	k=3	1.10	.20	.30
	k=4	0.87	.40	.40
	k=5	0.45	.60	.60

Source: Authors' calculations using Romer series.

TABLE 9
INTERNATIONAL OUTPUT GAPS

Period	United Kingdom Horizon			France Horizon			Germany Horizon		
	3	5	8	3	5	8	3	5	8
Post-WWII	.011	.016	.019	.007	.010	.012	.012	.016	.019
Pre-WWI	.019	.026	.030	.019	.026	.028	.011	.015	.019
Period	Sweden Horizon			Canada Horizon					
	3	5	8	3	5	8			
Post-WWII	.010	.015	.018	.012	.021	.025			
Pre-Depression	.022	.027	.031	.050	.057	.065			

Source: Authors' calculations using data from *Phases of Capitalist Development*.

TABLE 10
INTERNATIONAL BIAS-ADJUSTED VARIANCE RATIOS

Period	United Kingdom Horizon			France Horizon			Germany Horizon		
	2	5	10	2	5	10	2	5	10
Post-WWII	0.79	0.27	0.16	1.19	1.01	0.67	1.51	2.10	3.14

Pre-WWI	1.19	0.82	0.59	0.77	0.41	0.36	0.99	0.49	0.34
	Sweden			Canada					
	Horizon			Horizon					
Period	2	5	10	2	5	10			
Post-WWII	1.37	1.26	1.28	1.00	0.98	1.30			
Pre-Depression	1.06	1.17	0.96	0.79	0.37	0.28			

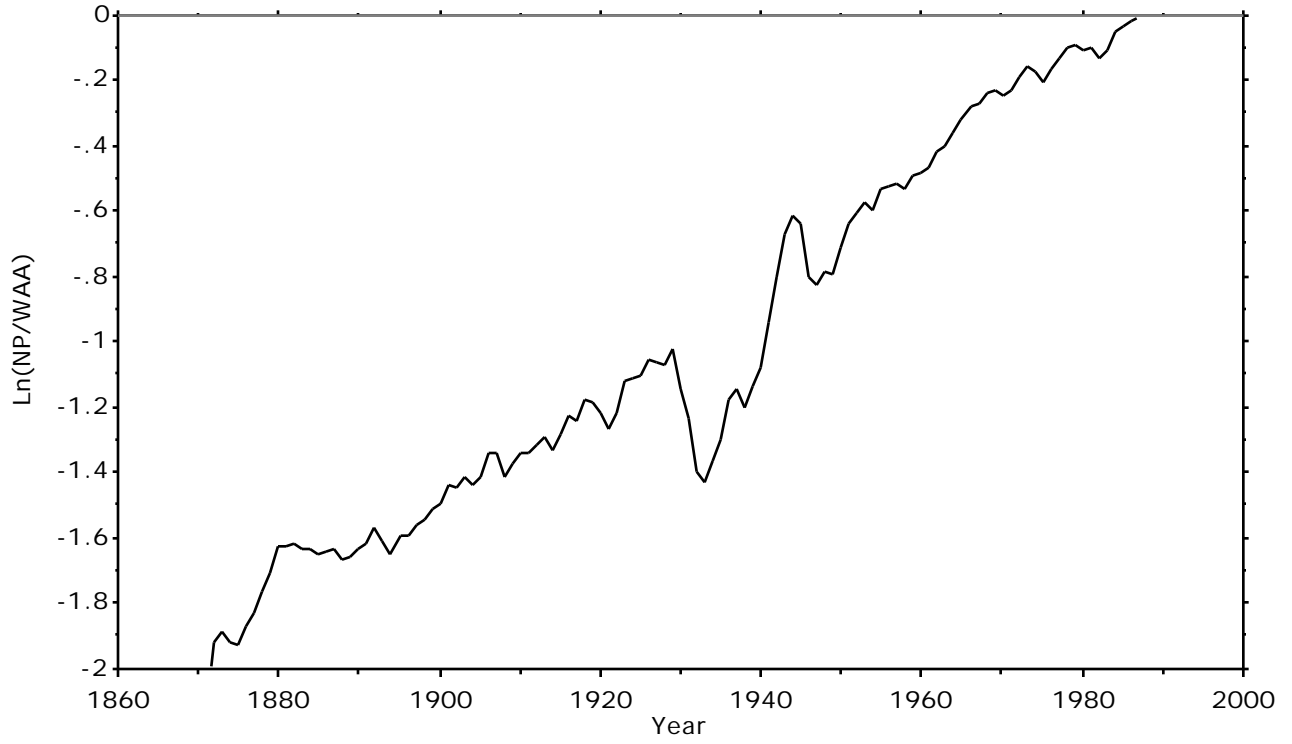
Source: Authors' calculations using data from *Phases of Capitalist Development*.

TABLE 11
PROBABILITY OF DEPRESSION AND RECOVERY IF OUTPUT DOES FOLLOW
A RANDOM WALK

Period from Which Innovation Variance Drawn:	Estimated Innovation Std. Dev.:	Probability of Recovery Given Decline:
1919-1939	7.1%	1.06%
1889-1933	5.3%	.04%
1889-1929	3.7%	2*10 ⁻⁵ %

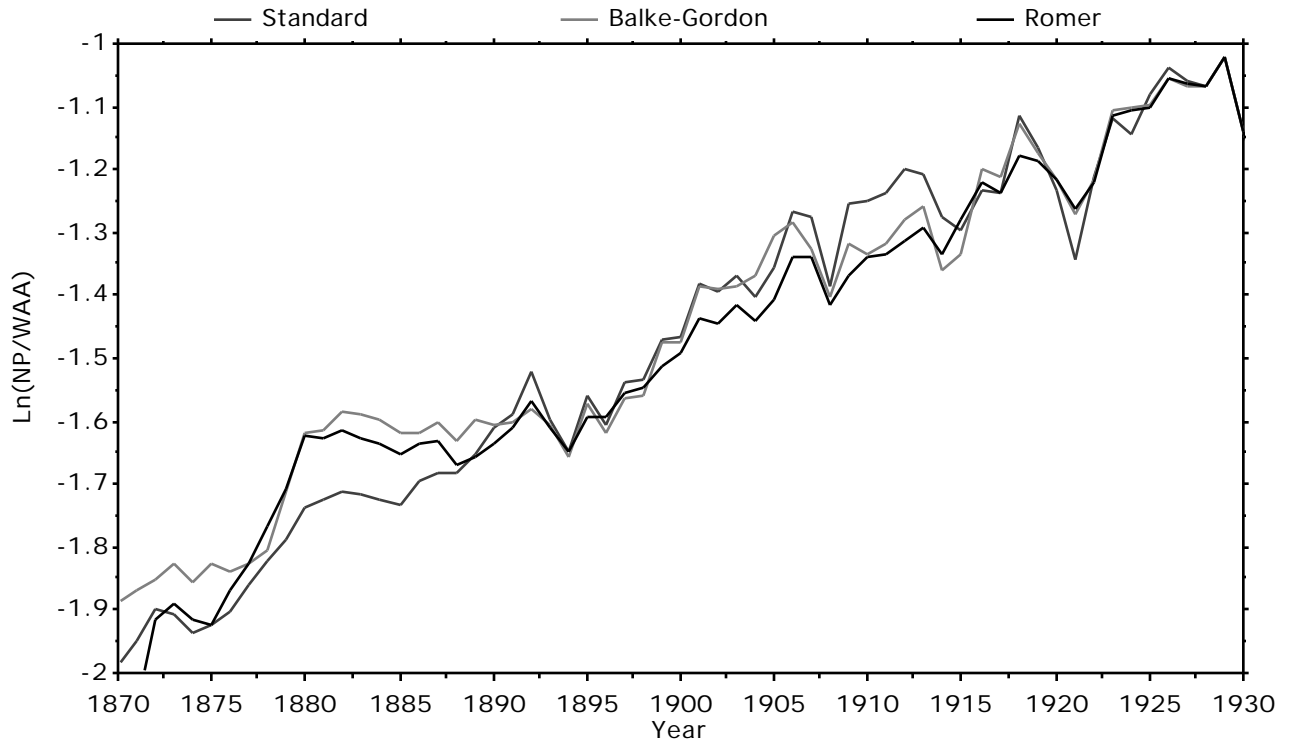
Source: Authors' calculations using Romer series.

FIGURE 1
ROMER ESTIMATES OF U.S. OUTPUT PER PERSON OF WORKING AGE, 1870-1987



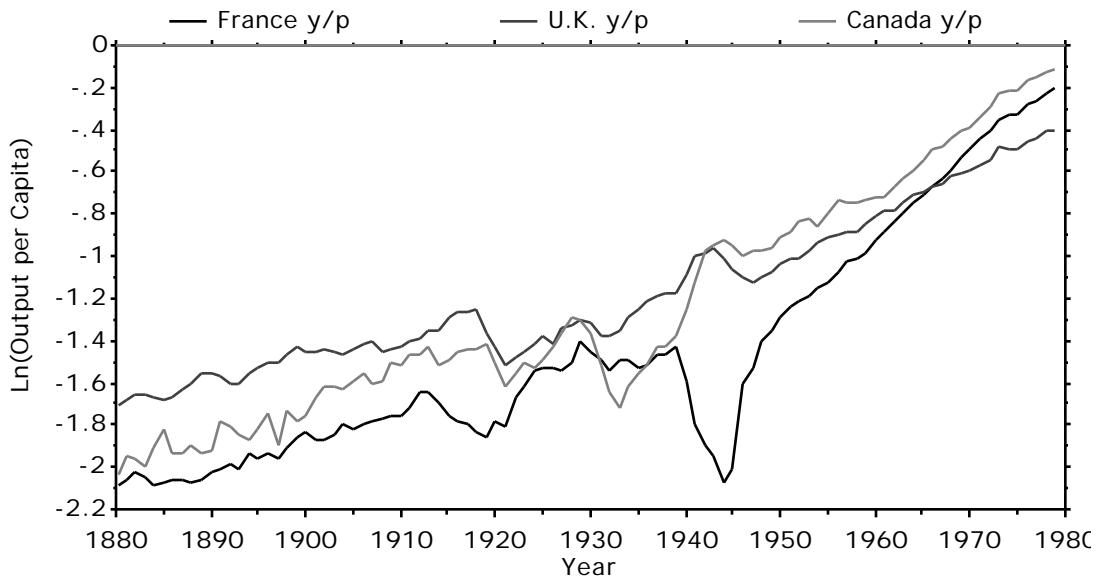
Source: Romer output series from "Prewar Business Cycle," and "GNP 1909-28," and *NIPA*.

FIGURE 2
STANDARD, ROMER, AND BALKE-GORDON ESTIMATES, 1870-1930

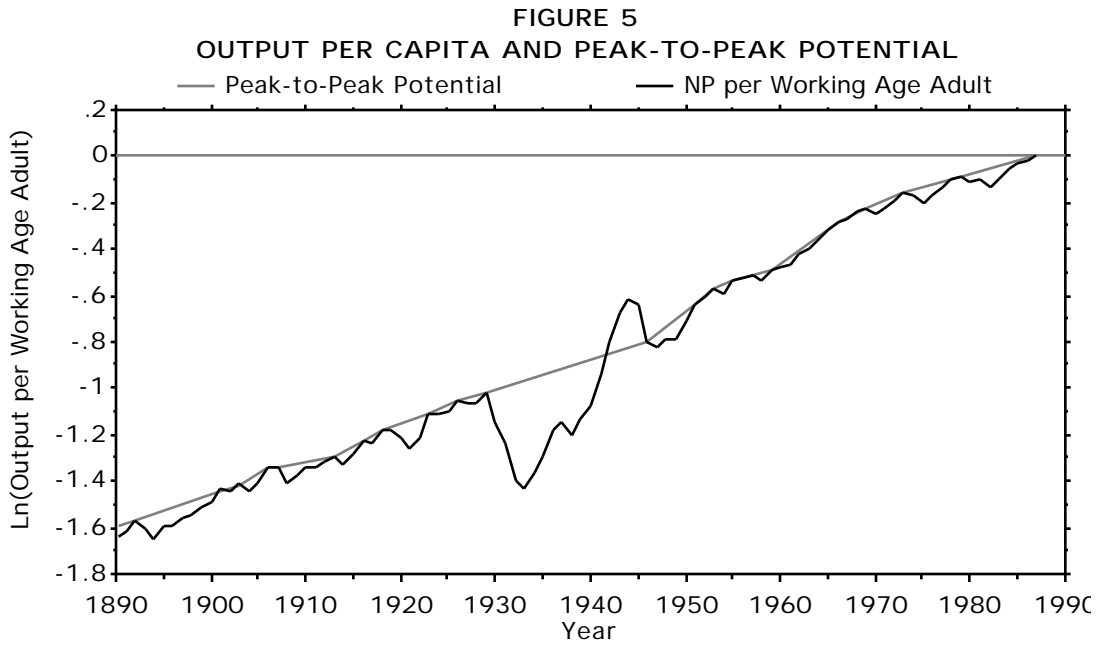
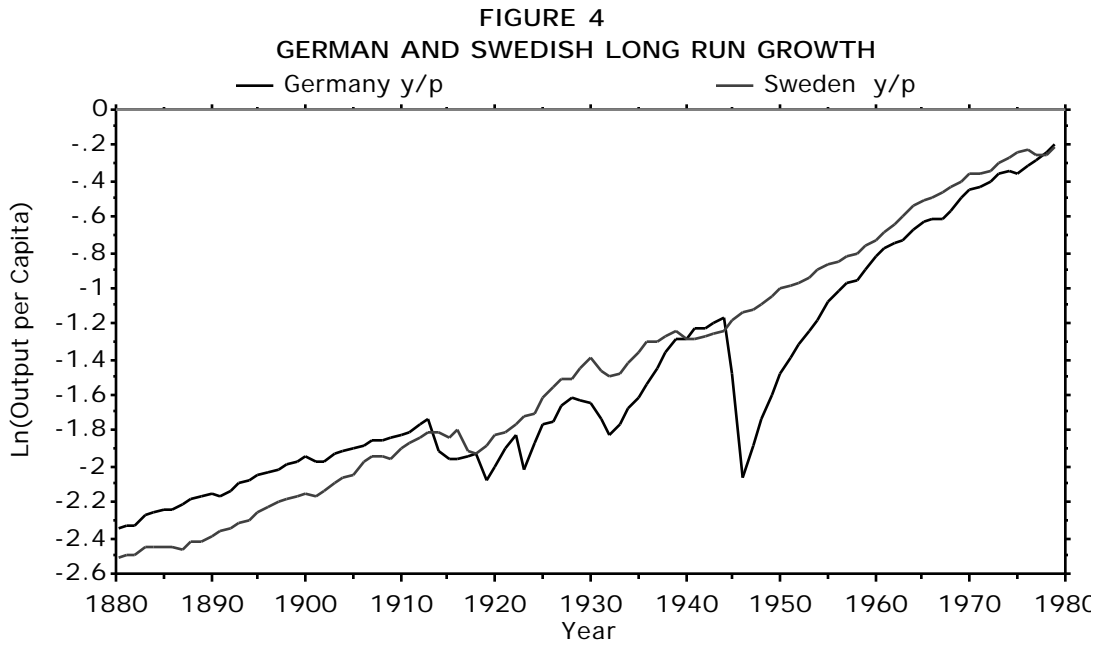


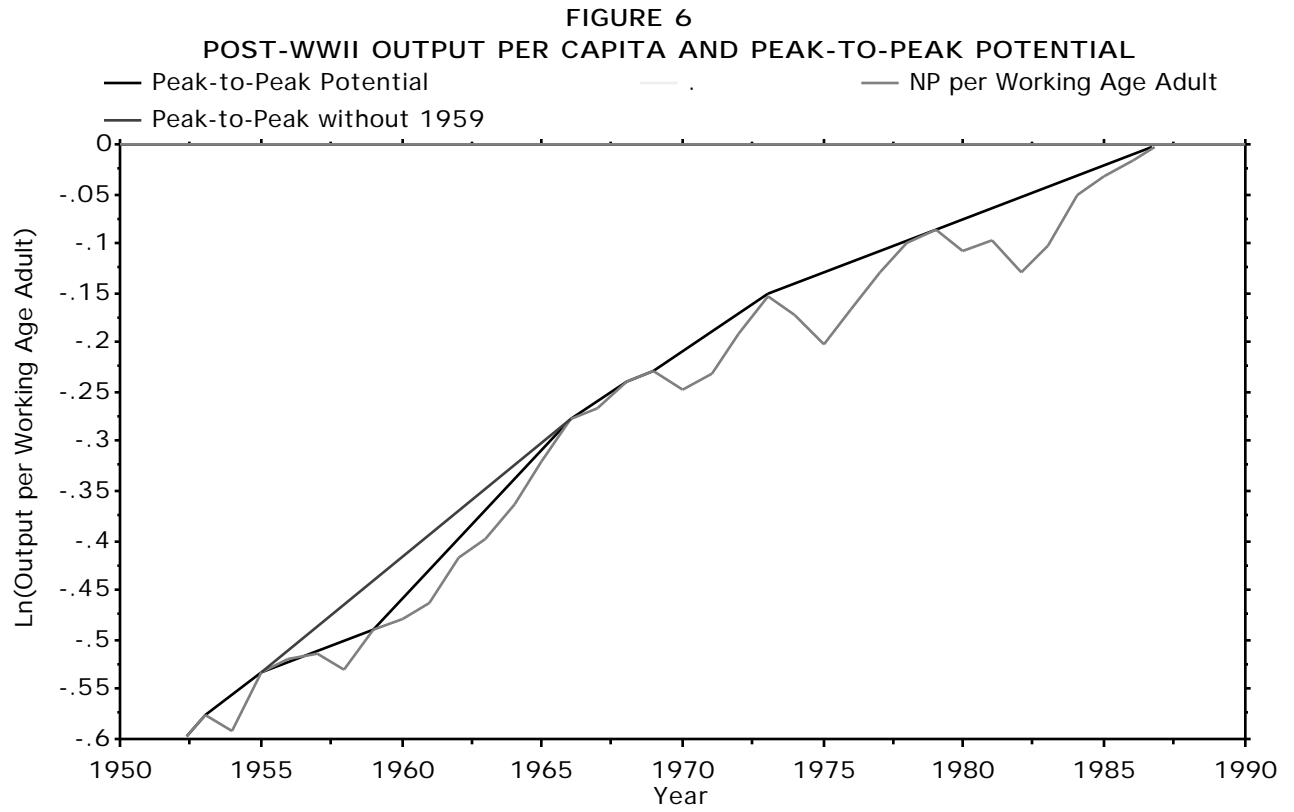
Source: As in table 1.

FIGURE 3
FRENCH, BRITISH, AND CANADIAN LONG RUN GROWTH



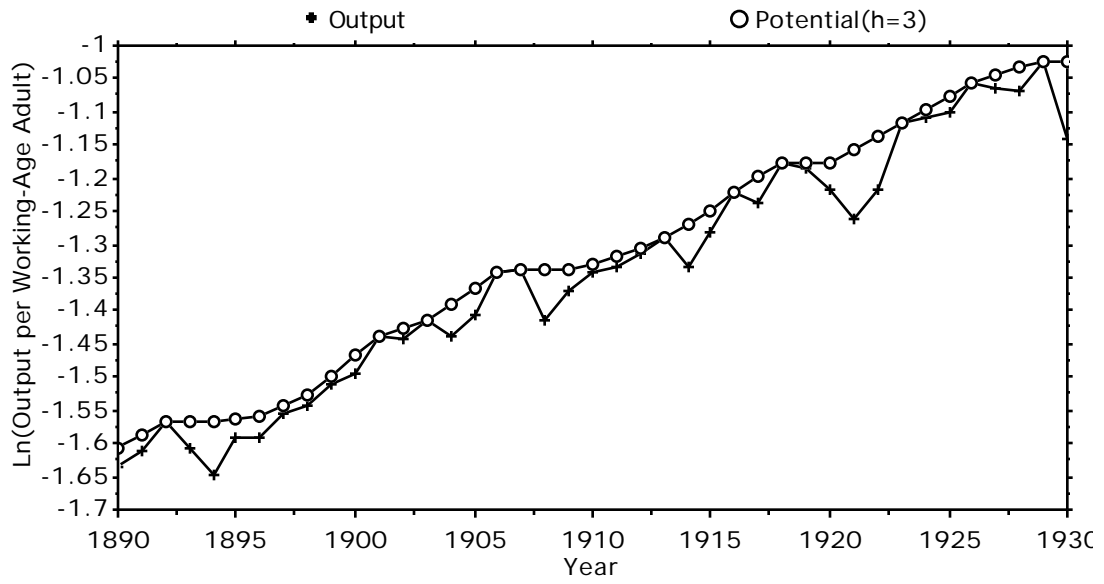
Source: Maddison, *Phases of Capitalist Development*.





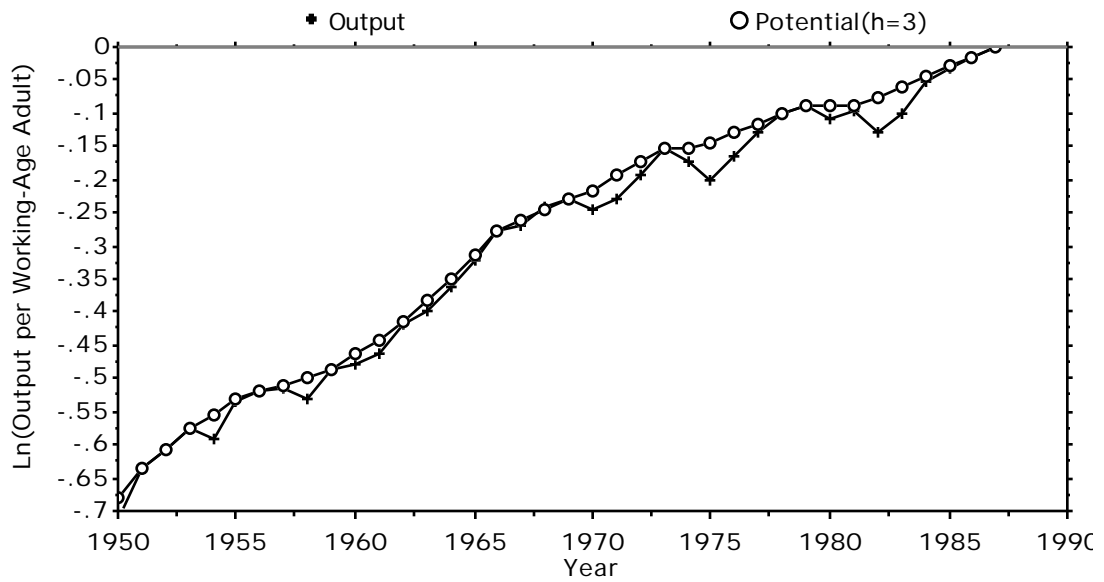
Source: Authors' calculations using *NIPA*.

FIGURE 7
PEAK-TO-PEAK POTENTIAL (3 YR HORIZON)



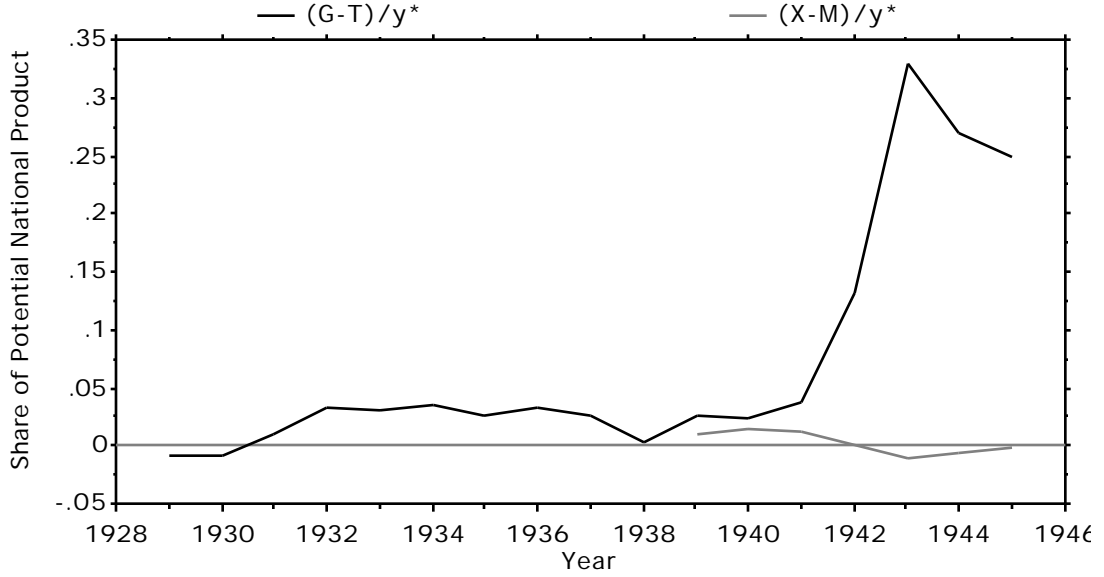
Source: Authors' calculations using the Romer series.

FIGURE 8
PEAK-TO-PEAK POTENTIAL (3 YR HORIZON)



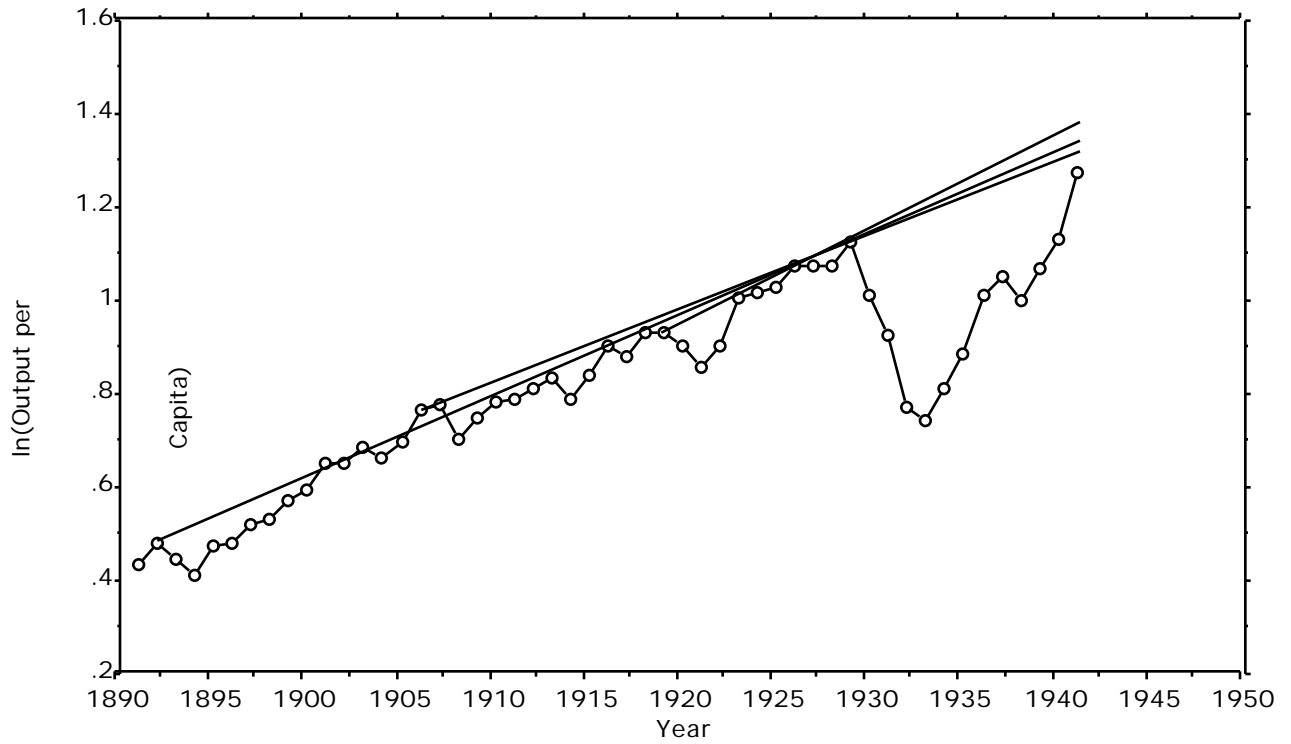
Source: Authors' calculations using the NIPA.

FIGURE 11
BUDGET AND TRADE BALANCE RELATIVE TO POTENTIAL NATIONAL PRODUCT



Source: NIPA.

FIGURE 12
RECOVERY FROM THE GREAT DEPRESSION FOR DIFFERENT TREND LINES



Source: Authors' calculations using Romer series.

NOTES