POTENTIAL GNP: ITS MEASUREMENT AND SIGNIFICANCE

Arthur M. Okun, Cowles Foundation, Yale University*

Potential GNP and Policy

"How much output can the economy produce under conditions of full employment?" The concept and measurement of potential GNP are addressed to this question. It is a question with policy significance because the pursuit of full employment (or "maximum employment" in the language of the Employment Act) is a goal of policy. And a target of full employment of labor needs to be linked to a corresponding target of full employment output, since policy measures designed to influence employment operate by affecting aggregate demand and production. How far we stand from the target of full employment output is important information in formulating fiscal and monetary policy. Thus, quantification of potential output offers one of the guides to stabilization policy and one indicator of its success.

The quantification of potential output -- and the accompanying measure of the "gap" between actual and potential -- is at best an uncertain estimate and not a firm, precise measure. While there are more precise measures of economic performance, they are not fully substitutable for the concept of potential output. To appraise the vigor of an expanding economy, it is important and enlightening to study customary cyclical measures, such as advance over previous peak levels or advance over recession trough levels. But these measures do not tell us how far we have to go to meet our targets, unless we are prepared to assume that each peak is like any other one and all troughs are likewise uniform. The record of the past decade testifies to the dramatic differences among cyclical peaks in levels of resource utilization.

The evaluation of potential output can also help to point up the enormous social cost of idle resources. If programs to lower unemployment from 5-1/2 to 4 percent of the labor are viewed as attempts to raise the economy's "grade" from 94-1/2 to 96, the case for them may not seem compelling. Focus on the "gap" helps to remind policy-makers of the large reward associated with such an improvement.

---

* My research in this area was done principally while I served on the Staff of the Council of Economic Advisers, and I had the benefit of many helpful comments and suggestions from members of the Council and the Staff. But the views reported here are my own and do not necessarily reflect those of the Council.

The Four Percent Unemployment Rate

Potential GNP is a supply concept, a measure of productive capacity. But it is not a measure of how much output could be generated by unlimited amounts of aggregate demand. The nation would probably be most productive in the short-run with inflationary pressure pushing the economy. But the social target of maximum production and employment is constrained by a social desire for price stability and free markets. The full employment goal must be understood as striving for maximum production without inflationary pressure; or, more precisely, as aiming for a point of balance between more output and greater stability, with appropriate regard for the social valuation of these two objectives.

It is interesting and perhaps surprising that there seems to be more agreement that a four percent unemployment rate is a reasonable target under existing labor market conditions than on any of the analytical steps needed to justify such a conclusion. Economists have never developed a clear criterion of tolerable price behavior or any quantitative balancing of conflicting objectives which could be invoked either to support or attack the target of a four percent rate. Indeed, I should expect that many economists who agree on the four percent target would disagree in estimating how prices and wages would behave if we were on target. Nor can the four percent rate be said to meet Beveridge's criterion for full employment -- that job vacancies should be equal to the number of unemployed. We simply have no count of job vacancies and could not possibly translate Beveridge's goal into any available measure of unemployment.

Having said what the four percent unemployment rate is not, I shall now state that it is the target rate of labor utilization underlying the calculation of potential GNP in this paper. The statistical and methodological problems would not be altered if a different rate were selected; only the numbers would be changed.

Potential GNP as a Short-run Concept

In estimating potential GNP, most of the facts about the economy are taken as they exist: technological knowledge, the capital stock, natural resources, the skill and education of the labor force are all data, rather than variables. Potential differs from actual only because the potential concept depends on the assumption -- normally contrary to fact -- that aggregate demand is exactly at the level that yields a rate of unemployment equal to four percent of the civilian labor force. If,

Reprinted from the 1962 Proceedings of the Business and Economic Statistics Section of the American Statistical Association

in fact, aggregate demand is lower, part of potential GNP is not produced; there is unrealized potential or a "gap" between actual and potential output.

The failure to use one year's potential fully can influence future potential GNP: to the extent that low utilization rates and accompanying low profits and personal incomes hold down investment in plant, equipment, research, housing, and education, the growth of potential GNP will be retarded. Because today's actual output influences tomorrow's productive capacity, success in the stabilization objective promotes more rapid economic growth.

The Measurement Problem

As it has been defined above, potential output is observed only when the unemployment rate is four percent, and even then must be viewed as subject to stochastic variation. At any other time, it must be regarded as a hypothetical magnitude. The observed actual measures of labor utilization tell us by a simple arithmetic calculation how much employment would have to increase, given the labor force, to make the unemployment rate four percent. But they do not offer similar direct information on other matters that might make labor input at full employment different from its observed level:

a) how average hours worked per man would be altered if the level of aggregate demand were consistent with full employment;

b) how participation rates in the labor force -- and hence the size of the labor force -- would be affected under conditions of full employment.

Nor do the actual data reveal directly what aggregate labor productivity would be under full employment conditions. There are many reasons why productivity might be altered in the aggregate; the added workers, changed average hours, possible alterations in the sectoral distribution of employment, higher utilization rate of capital, and altered efficiency in the use of employees all could make a difference in productivity at full employment.

The Leap from Unemployment to Output

Ideally, the measurement of potential output would appraise the various possible influences of high employment on labor input and productivity and evaluate the influences step-by-step, developing quantitative estimates for each adjustment to produce the desired measure of potential. While I shall discuss the steps individually below, the basic technique I am reporting consists of a leap from the unemployment rate to potential output rather than a series of steps involving the several underlying factors. Strictly speaking, the leap requires the assumption that, whatever the influence of slack economic activity on average hours, labor force participation, and manhour productivity, the magnitudes of all these effects are related to the unemployment rate. With this assumption, the unemployment rate can be viewed as a proxy variable for all the ways in which output is affected by idle resources. The measurement of potential output then is simplified into an estimate of how much output is depressed by unemployment in excess of four percent.

Statistical estimates

The answer I have to offer is simple and direct. In the postwar period, on the average, each extra percentage point in the unemployment rate above four percent has been associated with about a three percent decrement in real GNP. This result emerged from three methods of relating output to the unemployment rate.

1. -- First differences -- In one technique, quarterly changes in the unemployment rate (Y), expressed in percentage points, are related to quarterly percentage changes in real GNP (X). This regression equation, fitted to 55 quarterly observations from 1947-II to 1960-IV, yields:

$$Y = .30 - .30X \quad (r=.79)$$

According to this estimate, the unemployment rate will rise by 0.3 points from one quarter to the next if real GNP is unchanged, as secular gains in productivity and growth in the labor force push up the unemployment rate. For each extra one percent of GNP, unemployment is 0.3 points lower. At any point in time, taking previous quarters as given, one percentage point more in the unemployment rate means 3.3 percent less GNP.

2. -- Trial gaps -- A second method consists of selecting and testing certain exponential paths of potential output, using alternative assumed growth rates and benchmark levels. The percentage "gaps" implied by these paths are then related to the unemployment rate (U) using a regression equation: $U = a + b\,(gap)$. The criteria for judging the validity of the assumed potential paths are: 1) goodness of fit; 2) absence of any trend in the residuals; 3) agreement with the principle that potential GNP should equal actual GNP when $U = 4$.

The slope terms in this equation fitted to various paths and different periods consistently ran from .28 up to .38. One such equation was reported in the March 1961 statement of the Council of Economic Advisers to the Joint

Economic Committee. It was:

$$U = 3.72 + .36 \text{ gap } (r=.93)$$

where the gap was derived from a 3-1/2 percent trend line through actual real GNP in mid-1955. The equation was fitted to quarterly data for 1953-60. It implies that an increment of unemployment of one percent is associated with an output loss equal to 2.8 percent of potential output -- or a somewhat larger percentage of actual output when actual is below potential. The estimated unemployment rate associated with a zero gap is 3.72 percent, not too far from the 4.0 percent ideal.

3. -- Fitted trend and elasticity -- The first method described above relied on the use of changes in GNP and in unemployment. The second method used levels but assumed the trend of output-growth at constant unemployment rates. It is also possible to derive the output-unemployment coefficient from data on levels without assuming a trend. The following model permits such a calculation:

a. -- There is a constant elasticity relationship in the relevant range between the ratio of actual (A) to potential (P) output, on the one hand, and the "employment rate" (N=100-U) as a fraction of its potential level (N_F):

$$\frac{N}{N_F} = \left(\frac{A}{P}\right)^a$$

b. -- There is a constant growth rate (r) of potential output starting from some level P_o such that at any time t:

$$P_t = P_o \ e^{rt}$$

By substitution and rearrangement:

$$N_t = \frac{A_t^a \cdot N_F}{P_o^a \cdot e^{art}}$$

Logarithmically:

$$\log N_t = \log \frac{N_F}{P_o^a} + a \log A_t - (ar)\, t$$

The log of the employment rate is here related to a time-trend and to the log of actual real GNP. When a regression equation is fitted to log N as the dependent variable and log A and t as independent variables: 1.) -- The coefficient of log A is the "output elasticity of the employment rate;" 2.) -- The coefficient of time is the product of that elasticity and the potential growth rate; it therefore yields an estimate of the potential growth rate; and 3.) -- The intercept yields the benchmark (P_o) for any given N_F, here taken as 96.

Fitted to varying sample periods, the estimated elasticity coefficient ran .35 to .40, suggesting that each one percentage point reduction in unemployment means slightly less than a 3 percent increment in output (near the potential level). The trend growth rate, fitted to 1947-60 quarterly data, was 3.9 percent, but it was clear that this was not uniform throughout the period. For the post-Korean period, the estimated trend growth in potential was near 3-1/2 percent, while, for the 1947-53 period, it was near 4-1/2 percent.

The uniformity that emerged from these various techniques was the approximate 3-to-1 link between output and the unemployment rate. My own subjectively weighted average of the relevant coefficients is 3.2, yielding the following estimate of potential:

$$P = A\,[1 + .032\,(U - 4)]$$

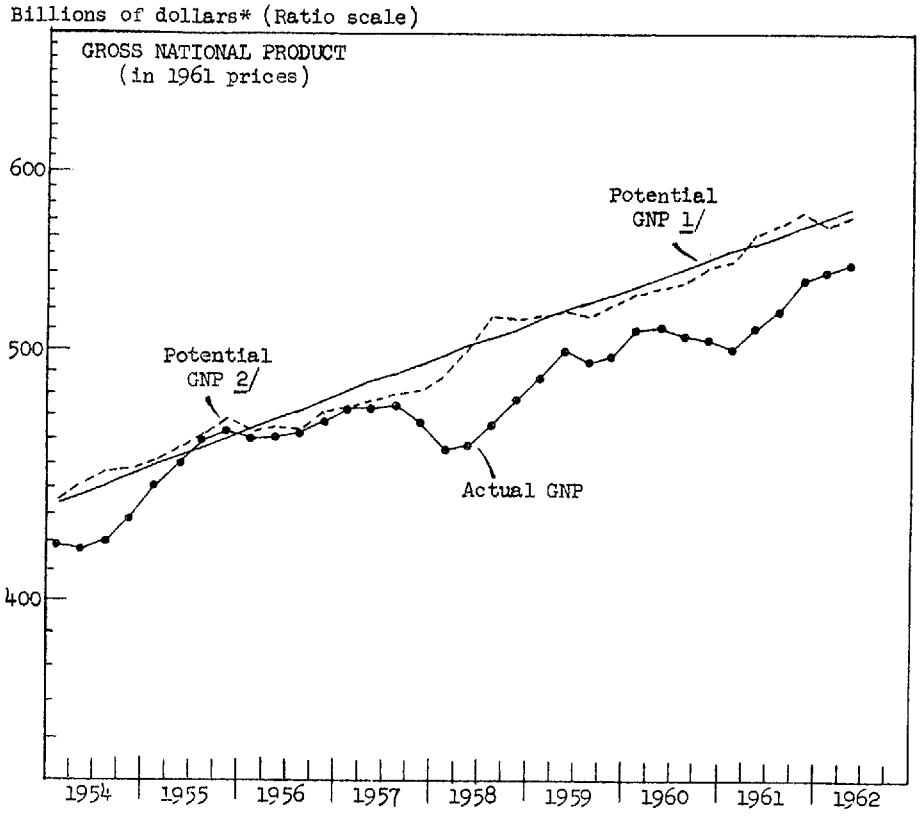
When the unemployment rate is four percent, potential GNP is estimated as equal to actual; at a five percent rate of unemployment, the estimated "gap" is 3.2 percent of GNP. In the periods from which this relationship was obtained the unemployment rate varied from about 3 to 7-1/2 percent; the relation is not meant to be extrapolated outside this range. I have no reason to expect the 3.2 coefficient to apply if unemployment were either 1 or 15 percent of the labor force.

Smoothing the Potential Path

The dashed line in the accompanying figure shows the implied time-series of potential GNP derived by applying the 3.2 coefficient to excess unemployment for the period 1954 to date. The result is a curve that wiggles from quarter to quarter, even dipping at times. The dips and small increases in estimated potential are concentrated in advanced stages of expansion -- 1956-57, 1959, and early 1962. Quarters of rapid rise in estimated potential output occur in early expansion -- 1955, 1958, 1961.

The question that arises is whether (1) these wiggles and jiggles should be taken seriously, as indications of irregular or cyclical patterns in the growth of productive capacity or (2) whether they should be attributed to an imperfect correlation of the unemployment rate with unused potential output. In the former case, the irregular path upward shown by the dashed line would be the estimated series of potential GNP. In the latter case, some

Potential and Actual GNP, 1954-1962



* Seasonally adjusted annual rates.
1/ 3-1/2% trend line through middle of 1955.
2/ Based on unemployment rate.

smoothing of that irregular path would be in order.

One way of smoothing which eliminates all the ripples is to substitute a simple exponential curve that corresponds with the trend and level of the wiggly series. Such a line is obtained by a trend that goes through actual output in mid-1955 as a benchmark and moves upward at a 3-1/2 percent annual rate. The trend measure of potential is shown as the solid line in the figure. It presents an opposite extreme alternative -- the view that the upward path of potential GNP has been perfectly smooth in the post-Korean period. On the whole, the two measures agree quite well. A trend line with either a 3 or a 4 percent growth rate -- or with a markedly different "benchmark" level -- would clearly not fit the dashed line equally well. In general, periods of early expansion -- like 1955; 1958-II to 1959-I; and 1961-II to 1961-IV -- show larger gaps by the unemployment measure than by the trend technique. The reverse is true for late expansion and recession periods, like 1956-II to 1958-I and 1959-III to 1961-I.

My own inclination is to select the smooth trend measure of potential output for the post-Korean period. I find it difficult to accept the verdict that potential output has actually contracted at times, as the unsmoothed unemployment measure implies. Nor can I believe

that the economy's productive capacity rises most rapidly in early expansion, even though actual production may be increasing briskly. This is not the period when investment expenditures -- much less completed investment projects' -- are at a peak; nor is it a time of heavy innovations, by an external evidence I know.

The spurts shown in early expansion periods can be accounted for by the hypothesis that unemployment lags somewhat behind the movement of output, and therefore is slow to decline in early recovery. Indeed, in statistical tests of some of the regression equations reported above, it was found that unemployment in the current quarter depends on past as well as current levels of GNP, with a higher level of past output meaning less current unemployment. This implies that decisions on hiring labor for next quarter are strengthened by a high level of current output.

The cyclical ripples in the unemployment measure may also reflect, in part, a lead of the workweek in advance of employment. Total manhours worked rise more rapidly than employment in early expansion and less rapidly in late expansion. The initial impact of a change in the pace of economic activity is particularly strong on the workweek and is later shifted more fully onto employment. Presumably, this lagged effect might be incorporated into the estimate of potential based on the unemployment rate, in such a way as to smooth that potential curve and bring it closer to the trend estimate of potential. But, for the post-Korean period, there is no obvious shift in the trend of potential; and the 3-1/2 percent trend line, while obviously too smooth a time path, fills the assignment rather well.

The trend estimate of potential for the 1954-62 period still rests on the unemployment-output relationship reviewed above, that an excess of 1 point in the unemployment rate means, on the average, a loss of about 3 percent in output. The trend line, however, suggests that the output loss per point of the unemployment rate exceeds 3 percent in late expansion and in recession and is somewhat less than 3 percent in early expansion.

It should be noted that this trend does not fit the earlier postwar years. If one projected the 3-1/2 percent trend back to 1947, the trend-technique would clearly overestimate potential output. The indicated potential growth of the 1947-53 period is nearer to 4-1/2 percent. The lower potential growth rate of the post-Korean period is associated, in part, with less success in making full use of our potential. The "gaps" between potential and actual have held down the size and held up the average age of our capital stock, thereby lowering the growth of potential.

The Steps

The findings above assert that a reduction in unemployment, measured as a percentage of the labor force, has a much larger than proportionate effect on output. To appraise and evaluate this finding, it is necessary to inspect the steps which were leaped over in the statistical relationships between output and unemployment. Clearly, the simple addition of one percent of a given labor force to the ranks of the employed would increase employment by only slightly more than one percent:

$\frac{100}{100-U}$ percent to be exact. If the workweek and productivity were unchanged, the increment to output would be only that 1 + percent. The 3 percent result implies that considerable output gains in a period of rising utilization rates must stem from some or all of the following: induced increases in the size of the labor force; longer average weekly hours; and greater productivity.

Labor force

Participation in the labor force as we measure it consists of either having a job or seeking actively to work. The resulting measures of labor force are not pure reflections of supply; they are affected by job availability. In a slack labor market, people without a job may give up when they are convinced that job-hunting is a hopeless pursuit. They then may be viewed as having left the labor force though they stand ready and eager to work. Furthermore, there are secondary or passive members of the labor force who will not actively seek employment but would accept gainful employment if a job came looking for them. This latter group suffers little or no personal hardship in not having work, but the output they would contribute in a fully employed economy is a relevant part of the nation's potential GNP.

There may be induced changes in the labor force in the opposite direction: e.g., the loss of a job by the breadwinner of a family might increase the measured labor force by leading his wife and teen-age children to seek work. The prewar literature debated the probably net effects of these opposing influences on participation rates. However, the postwar record has convincingly delivered the verdict that a weak labor market depresses the size of the labor force. But the magnitude and timing of the effect is not clear.

Even the conceptual problem of defining a potential labor force is difficult -- we should not wish to count only the secondary labor

force members who would appear for work tomorrow morning; on the other hand, we would not want to include all those who might be attracted by many years of continued job availability. The response of participation rates is likely to be a complicated lagged phenomenon which will not be closely tied to the current unemployment rate. While this aspect of the difference between potential and actual output is hard to quantify, zero is certainly not a satisfactory estimate. At the end of 1960, the Bureau of Labor Statistics estimated the difference between actual and "normal" labor force at 561,000. If this figure is taken as the induced effect of poor opportunities for jobs, it implies that, in those recession conditions, for every 10 people listed as unemployed over and above the 4 percent rate, there were three additional potential workers who were not actively seeking work.

Hours

Taking into account the normal secular decline in hours worked per man, there is a clear relationship between movements in average hours and in output. When output has been rising rapidly, average hours have expanded -- or, at least, have not contracted. On the other hand, periods of low growth or decline in GNP mean more rapid declines in average hours per man. The data point toward the concept of a full employment path of average annual hours. But the concept of full employment hours is hard to quantify: e.g., in a rapid rise of output toward full employment, the amount of overtime might well push the workweek above the level consistent with steady full employment. Furthermore, economy-wide data on average hours are notoriously poor. However, using what evidence is available, we find that each one percent difference in output is associated with a difference of 0.14 percent in hours per man, including both overtime and part-time work.

The figure of 0.14 is obtained by fitting a least-squares regression line to annual data for 1947-59. The data are found in the Bureau of Labor Statistics Release (USDL-4155) of June 28, 1960. The variables are percent change in manhours of work per person employed (Y) and percent change in private nonagricultural output (X), restricted to private nonagricultural output and employment; establishment figures are the source of the man-hour estimates. The fitted line is:

$$Y = 0.843 + .142X \quad (r = .85)$$

When this equation is used to compare average hours for different possible outputs at the same point in time, the 0.142 coefficient reflects the percentage difference in hours per man that accompanies a one percent difference in output.

Returning to the finding that a one percentage point reduction in the unemployment rate means 3.2 percent more GNP, the hours-output estimate above indicates that it will also be accompanied by an increase of nearly one half of one percent in hours per man, or an addition of about 0.2 of an hour to the workweek. With an allowance for induced gains in labor force, based illustratively on the 1960 estimate cited above, the reduction of one point in the unemployment rate means perhaps a 1.8 percent increase in total labor input measured in manhours. Then, to get the 3.2 percent increment in output, manhour productivity must rise by about 1.4 percent.

Productivity

The direct checks that could be made on productivity data were consistent with this implication of the output-unemployment relationship. The record clearly shows that manhour productivity is depressed by low levels of utilization, and that periods of movement toward full employment yield considerably above-average productivity gains.

The implications and explanations of this phenomenon are intriguing. Indeed, many a priori arguments have been made for the reverse view -- that depressed levels of activity will stimulate productivity through pressure on management to cut costs, through a weeding-out of inefficient firms and low quality workers, and through availability of more and higher quality capital per worker for those employees who retain their jobs. If such effects exist, the empirical record demonstrates that they are swamped by other forces working in the opposite direction.

I have little direct evidence to offer on the mechanism by which low levels of utilization depress productivity. I can offer some speculation and try to encourage other researchers to pursue this problem with concrete evidence at a micro-economic level. The positive relationship between output and labor productivity suggests that much of labor input is essentially a fixed cost for fairly substantial periods. Thus high output levels permit the spreading of labor overheads, and low production levels raise unit fixed costs of labor. At times, we may take too seriously our textbook examples which view labor as a variable factor, with only capital costs as fixed. Even the most casual empiricism points to an overhead component in labor costs. There are many reasons why employment may not be easily variable:

1. -- Contractual commitments may tie the hand of management in a downward direction -- employees may have guaranteed annual wages,

supplementary unemployment compensation, rights to severance pay, etc. as well as actual contracts for a term of employment.

2. -- Technological factors, in a broad sense, may also be important. A firm plans on a division of labor and degree of specialization attuned to "normal" operations. If operations fall below normal, there may be marked indivisibilities which prevent the firm from curtailing its employment of specialists, clerical and sales personnel, and supervisors in parallel with its cutback in output.

3. -- Transactions costs associated with laying off labor and then, in the future, doing new hiring may be another influence retarding the adjustment of labor input to fluctuations in sales and output.

4. -- Acquired skills that existing employees have learned on the job may make them particularly valuable to the firm so that it pays to stockpile underemployed labor rather than run the risk of having to hire untrained men when business conditions improve.

5. -- Morale factors may also make layoffs undersirable.

All of these factors could help explain why slack economic activity is accompanied by "on-the-job underemployment," reflected in depressed levels of manhour productivity. Firms obviously do lay off labor in recession but they do so reluctantly. Their problems may be mitigated, in part, by the presence of voluntary quits which permit a downward adjustment of employment without layoffs. In part, the impact of slack on manhour productivity may be reduced by shortening average hours to spread the work and the wage-bill without a cut in employment. But these appear to be only partial offsets.

To the extent that the productivity losses of recessions are associated with fixity of labor costs, they would not be maintained indefinitely. If the recession was of long duration -- or merely was expected to last a long time -- firms would adjust their employment more drastically. On this reasoning, in an era when business cycle dips are continually short and mild, one might expect productivity to bear more of the brunt of recession and labor input to be less affected, even relative to the decline in output.

Changes in the level of economic activity are associated with shifts in the composition of employment and output by industry. A slack economy is accompanied by particularly depressed output in durable-goods manufacturing industries, where output per manhour is especially high. My own intuition suggested that this might be an important explanation of the relationship between productivity and the unemployment rate. But calculations on the change in composition from recession to recovery years indicate that, while shifts in industrial composition do influence aggregate productivity in the expected direction, the magnitude of the effect is trivial. There is some significance to the compositional shift between agriculture and nonagricultural industries. Manhour input in agriculture seems to be independent of overall economic activity in the short run, so all variations in labor input can be regarded as occurring in the nonagricultural sector. I assumed illustratively above that a point reduction in the unemployment rate means an increase in total manhours of 1.8 percent. If all of that 1.8 percent goes into nonagriculture, this would add 0.1 percent to economy-wide productivity (for given levels of productivity in each sector). This is still only a minor part of the total productivity gain that accompanies reduced unemployment.

Thus far, I have ignored the dependence of labor productivity on plant and equipment capacity. The entire discussion of potential output in this paper has, in effect, assumed that idle labor is a satisfactory measure of all idle resources. In fact, measures of excess capacity in industrial plant and equipment do show a close relationship to unemployment -- idle men are accompanied by idle machines. But the correlation is not perfect and operating rates in industry should be considered along with employment data as an indicator of the gap between potential and actual output. Obviously, if capital were fully employed while there was much unemployed labor, this would hold down the productivity gains that could be obtained through full employment of labor. Robert Solow did use capital stock data together with unemployment data in fitting a production function for 1929 to date (see the American Economic Review of May 1962). His estimates of potential output for the post-Korean period agreed remarkably well with those I am reporting.

Still, I shall feel much more satisfied with the estimation of potential output when our data and our analysis have advanced to the point where the estimation can proceed step-by-step and where the capital factor can be explicitly taken into account. Meanwhile, the measure of potential must be used with care. The trend line yields a point-estimate of the "gap," e. g., $31.3 billion for 1962-II. But that specific figure must be understood as the center of a range of plausible estimates. By my personal evaluation of its degree of accuracy, I find potential output useful -- and superior to substitute concepts -- for many analytical purposes.