

Capacity Utilization, Inflation, and
Aggregate Demand Policy

John Bryant and Tom Supel

June 1978

Working Paper #: 116

PACS File #: 1950

NOT FOR DISTRIBUTION

The views expressed herein are solely those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. The material contained is of a preliminary nature, is circulated to stimulate discussion, and is not to be quoted without permission of the authors.

Suppose there is excess plant capacity. Will a more stimulative monetary policy raise the rate of output growth and lower the unemployment rate without substantially raising the pace of inflation? The answer commonly given, particularly by policy makers and the journalists who report on policy, is "yes." However, many economists believe that "no" is the correct answer. This purpose of this paper is to discuss both replies in an attempt to clarify the issues. In fairness it must be stated at the outset that the authors belong to the "no" camp.

The paper is divided into three sections. In the first section the concept of capacity utilization is discussed. The second section presents the implications of capacity utilization for policy in alternatives models of the aggregate economy. The paper concludes with some empirical results bearing on the issues raised.

I. The Concept of Capacity Utilization.

Two definitions of capacity are examined: engineering capacity and economic capacity as defined by the Federal Reserve Bank of St. Louis.

A. Engineering Capacity.

For the case of a single machine "capacity" can frequently be given an unambiguous definition. Capacity is the maximal amount of output that can be produced by the machine with other inputs supplied in limitless quantities. The extension to the economy as a whole is that capacity is the amount of output with every machine working at capacity. This is not, of course, an operational definition. Raw material, labor, and land inputs are not available in limitless quantities, nor is there any guarantee that capacity output of intermediate good industries just matches input needs of capacity producing finished goods industries, nor can all technology easily be described by a capacity figure. This

definition of capacity does not have the subtleties of tradeoffs and substitutions summarized by a cost function.

Engineering capacity can, however, be given an interpretation in terms of cost curves. Assume that for the individual firm inputs, other than capital, are infinitely elastically supplied, the firm is a price taker. Assume further that for each identical machine the firm's output increases proportionately with inputs until capacity is reached. Then the firm's average variable cost function is flat up to capacity, and verticle thereafter. Capacity is the elbow in the cost curve. Can this definition of capacity be extended to the economy as a whole? Assume all firms are identical, and that inputs are costless shifted between firms. Even this is not enough to generate an elbow in the aggregate cost function. Inputs are assumed to be infinitely elastically supplied to the firm because the firm is small. However, all firms together are not small, so that in the aggregate inputs are not infinitely elastically supplied. And, of course, firms are not identical and inputs not costlessly shifted between them. The aggregate average cost curve, if it can indeed be defined, will not have an elbow. We see, then, that capacity has a very fuzzy meaning for the economy as a whole. The aggregate average cost curve may be relatively flat at low output and relatively steep at high output, but what point is capacity?

Presumably by capacity we mean some point where the average cost curve, and therefore the marginal cost curve and supply curve, is "fairly steep," and we will take this to be our operating definition of capacity.

Implicit in our discussion of capacity has been the assumption that there is no market for existing capital. Once a firm puts capital in place it can only be used by that firm for its designated use. This

may be a realistic assumption, particularly in the short run. But if, contrary to this assumption, there is a market in capital and firms are not identical, machines will not sit idle. A firm not using its capital will rent it out for another use. This too could introduce curvature into the cost function.

Given that capacity is such a fuzzy concept, why does it generate so much interest? One major school of economic thought believes that one of the assumptions behind the concept of capacity does hold in reality. As a result the average cost curve is fairly flat up to a point and fairly steep thereafter; while it is not an "elbow," the "elbow" is not a bad approximation. Specifically, Keynesians believe in multiple equilibria, including unemployment equilibrium, at a constant nominal wage. In unemployment equilibrium one major input, labor input, is infinitely elastically supplied in the aggregate up to a point. As long as the full employment level exceeds the point of output where capital becomes scarce, the average cost curve is flat up to that point and steep beyond it. One difficulty with this approach is that it still depends upon capital becoming scarce at once, something we have argued against above. This position would seem to argue instead for a smoothly rising average cost curve to the level where full employment is achieved, and steeper thereafter. Nevertheless, with this interpretation, a low rate of capacity utilization and a high rate of unemployment together can be taken as evidence of being on a relatively flat portion of the aggregate supply curve, and the reverse as evidence of being on a relatively steep portion.

B. Economic Capacity.

Because of the fuzziness of the concept of engineering capacity when applied to the aggregate economy, economists have tried to define

another concept of capacity. In this subsection we examine the definition used by the Federal Reserve Bank of St. Louis. This definition is:

"For any amount of fixed resources only one output rate can be produced using an efficient long-run model. This output rate is the economic capacity of the firm."

There are several problems with this definition. Suppose they mean the minimum of the long-run average cost curve. This is a funny definition of capacity because it is the output amount after firms have adjusted their capital stock, but capacity is generally assumed to be measured from given existing capital stock. It also is not a usable measure for the aggregate economy as it does not say how many firms are operating at capacity output. Further there is some good evidence that long-run average cost curves are flat bottomed, β capacity is not unique. Suppose capacity means that all existing firms are operating at the minimum of their short-run average cost curves (SAC). This may be a definable number but that it is useful for anything is in doubt. There is no presumption that currently existing firms should be at the minimum of this short-run average cost curves. Moreover, it is unclear what, if anything, the knowledge that the economy is at capacity thusly defined would tell you about the impact of a stimulative demand policy.

Defining capacity in terms of some sort of minimum cost is intuitively appealing and not unprecedented in the economics literature. However, we have great difficulty seeing the value in this new bit of jargon. The supply curve of the firm is still the relevant part of the marginal cost curve, and the aggregate supply curve is still the relevant summation. St. Louis is well aware of the relation of capacity to supply: "Aggregate capacity is only one point of the short-run aggregate

supply curve of the economy." In order to get any policy implications out of the concept of capacity, they go on to invoke the traditional Keynesian assumptions about labor supply and of a fixed capital stock. To produce an "elbow" (a vertical portion) in the supply curve, contrary to their previous assumptions of U-shaped SACs.

The principle conclusion of the section is that the concept of economic capacity is so nebulous that it does more harm than good in economic debates. The concepts of supply and demand are older but better. In the next section we turn to implications of capacity utilization for policy. In this discussion a high rate of utilization is interpreted as being on a relatively steep portion of the aggregate supply schedule, and a low rate of utilization is interpreted as being on a relatively flat portion. This may be a too generous interpretation of capacity utilization.

II. Policy Implications.

We now examine the policy implications of capacity utilization in two frameworks: (a) the Keynesian multiple equilibrium model and (b) the classical unique equilibrium model.

To recapitulate, Keynes argued that the economy is characterized by multiple output equilibria at different levels of demand, with low levels of demand corresponding to unemployment equilibria. To get this result Keynesians have assumed a rigid nominal wage, that labor is infinitely elastically supplied up to the full employment level. In contrast the classical assumption is that there is a unique (full employment) equilibrium.

For the sake of simplicity, we will discuss changes in the price level, rather than changes in the rate of inflation. First let us

consider a stimulative monetary policy in the Keynesian case, ignoring the possibility of a liquidity trap. Suppose there is unemployment and low-capacity utilization so that the supply curve is relatively flat. The increased money supply shifts the aggregate demand schedule higher, thereby raising the price level. With higher prices, producers find it profitable to hire more workers which they can do at the existing nominal wage as long as there is unemployment. Employment and output expand, and this expansion may very well be possible with only a modest increase in price if the amount of excess capacity is substantial. The distribution of the increase in nominal GNP between real output and price is largely determined by the slope of the supply curve. The increased prices may cause the price of some inputs to go up, shifting the supply curve up. However, this shift is likely to be small as the major variable input, labor, does not react to the higher prices. In short, a "permanent" gain in employment and output with possibly only a small increase in price can be obtained with an increase in the money supply. If there is unemployment, but capacity utilization is high, there will be a relatively small increase in output but a large increase in price. Finally, if there is full employment, output gains will only be temporary as the price increase causes the supply curve to shift up proportionately as in the classical case discussed below.

The "classical" alternative to the simple Keynesian case that we wish to consider assumes the following. First, the labor market is in equilibrium with workers satisfied with the real wage they are currently receiving. Second, excess plant capacity exists. Third, workers and labor leaders expect the current price level to continue. In this situation, an increase in the nominal wage appears to be an increase in the real wage and calls forth additional labor.

Can monetary policy have a real impact in this situation?

Consider an increase in money that lifts aggregate demand and the price level. Employers, because of higher prices, bid for more labor, thus raising the nominal wage. Employment and output expand. Sooner or later, however, labor realizes that the price level is higher and the real wage lower than anticipated. Price expectations adjust and higher nominal wages are demanded. The aggregate supply curve, in effect, shifts up at this point: the same output can only be supplied at a higher price level. Because of this shift, prices rise and output drops from its previous peak. But now there is a new and higher price level, and labor must a second time adjust its concept of the going price level. Again there is an upward shift in aggregate supply. The process continues until there is equilibrium in the labor market with actual prices equal to expected prices. Assuming the economy is homogenous in money and prices, output is back at its initial level, but the price level is higher by an amount that is proportional to the increase in the money supply.

Temporarily, there is an increase in output above its starting level, an increase that comes about because labor mistakenly believes that higher nominal wages mean higher real wages. But when workers correct their mistake and attempt to gain higher real wages, they eliminate the profit incentive that initially induced employers to expand. As workers are making very costly errors, this correction may take place rapidly. The final result is no gain in output. Nominal wages, prices, and profits are higher, but real wages and profits are unchanged.

The slope of the aggregate supply function is an element affecting the magnitude of the short-run trade-off. However, it is only one element, as the rate at which the supply curve shifts is at least as

important. Indeed the rate of shift is a major element in the argument of many Keynesians. Many would argue that the world is indeed classical in the long run, but the supply curve shifts only very slowly. Gains in output are indeed only temporary and bought at the cost of a long-run price rise, but if "temporary" means several years its well worth it. And if the supply curve can be treated as fixed in the short run its slope, as measured by the capacity utilization, determines the short-run trade off. However, it is not clear that a slowly shifting supply curve is any easier to explain than a fixed nominal wage. Nor is it clear how you are making people better off by stimulating output by fooling them into costly errors. Presumably the government is at best offsetting an error in the other direction imposed by an unpredictable economy.

One argument against the "classical" view is that it is obviously contradicted by experience.

Most usually, the evidence pointed to is the unemployment rate--it's too high. Our position is that the unemployment rate, tells us nothing about the validity of the "classical" view. Why? Well, on the one hand, because there is frictional unemployment, that is that natural rate of unemployment is not zero. What is important is the excess of unemployment above the natural rate. Moreover, even if there is excess unemployment, this does not imply that there is involuntary unemployment which more stimulative monetary or fiscal policy can cure. It may be that the job seekers' "reservation real wage" is simply above the market clearing real wage.

The fact that information is not perfect and that economic agents may be wrong in their expectations indicates that, there can be a short-run trade-off between output and prices: temporary gains in output can be achieved through unanticipated expansionary monetary

policy. However, these gains in output result from costly errors by the economic agents who are fooled. In addition, we pay for these temporary gains in output through a higher price level (inflation rate).

Excess capacity may be one of the ingredients of the shape of the short-run aggregate supply curve. For these reasons, it plays a role in the short-run dynamics of the economy. But this does not make it a fundamental factor in determining inflation. Indeed it is not.

III. Empirical aspects of capacity and inflation.

We have argued that, at best, the rate of capacity utilization provides information about the slope of the aggregate supply schedule. In this section we will examine empirical evidence bearing on this proposition. The following tests are joint hypotheses about the usefulness of capacity utilization and the validity of the Keynesian view of inflation. They say nothing about the validity of the "classical view."

In formulating our statistical tests, we will therefore assume that when the rate of capacity utilization is "low," the aggregate supply curve is rather flat, and when the rate of capacity utilization is high, the aggregate supply schedule is rather steep. We will further assume that price changes are due primarily to shifts in aggregate demand. In our judgment, this set of assumptions is consistent with traditional economic thinking over the past two or three decades. We comment below on the possibility that our statistical results are confounded by shifts of the aggregate supply that are induced by events such as the OPEC action.

This view of the world implies that there should be some significant positive correlation between price changes and the rate of

capacity utilization. When the economy is operating on the flat portion of the supply schedule, changes in aggregate demand will lead to changes in the rate of capacity utilization but will have only small price effects. At worst, the correlation between price changes and the rate of capacity utilization should be zero. When the economy is operating on the steep portion of the aggregate supply curve, changes in aggregate demand will produce large price changes with only modest changes in the rate of capacity utilization.

The rate of capacity utilization is only one of many variables that might explain inflation. In testing for the asserted positive correlation between price changes and capacity utilization, we must also try to account for other factors. That is, we need a model. But a reduced form model is adequate since we are only interested in whether or not the rate of capacity utilization helps to explain inflation. We therefore adopt the vector autoregression model as the primary framework for our statistical tests. In addition to representing an unconstrained reduced form, the vector autoregression is a legitimate representation of a wide class of stationary stochastic processes. Operationally, a vector autoregression is an ordinary type of regression with only lagged values of all model variables on the right-hand side.

The simplest possible model that meets our needs is a two variable model containing the rate of inflation and the rate of capacity utilization. The vector autoregressive representative of this model that we used is:

$$(1) \quad p_t = a + bT_t + \sum_{i=1}^8 c_i p_{t-i} + \sum_{i=1}^8 d_i C_{t-i} + \epsilon_t$$

$$(2) \quad C_t = e + fT_t + \sum_{i=1}^8 g_i p_{t-i} + \sum_{i=1}^8 h_i C_{t-i} + \delta_t$$

where

p = percent change in CPI.

C = FRB capacity utilization rate.

T = time.

This model was estimated with quarterly data for the period 1950.I to 1977.IV. Testing if the rate of capacity utilization helps to explain the rate of inflation is equivalent to testing the null hypothesis that the set of d_i 's in equation (1) are zero. This hypothesis cannot be rejected at significance levels below 10 percent.

A more complex model of the economy would contain more variables. For this purpose we used, in addition to p and C, the rate of change in real GNP, M1, and real business fixed investment, plus the unemployment rate and the 90-day Treasury bill rate. Once again the hypothesis that C does not help explain p cannot be rejected at significance levels below 10 percent.

These experiments provide, at best, weak support for the view that the capacity utilization rate helps to explain inflation. We did find that C entered the p equation in a statistically significant way when we tried the above complex model without the BFI variable. We have no satisfactory economic explanation of this since neither C nor BFI enters the price equation of the complex model. Note that this set of tests requires the assumption that prices tend to be inflexible downward so that high capacity is not associated with deflation.

It is possible that shifts in the supply schedule could be causing price increases and thus reducing the explanatory power of capacity. Since the OPEC shock of 1973 was the most prominent example of an unanticipated supply shock, the rate of capacity utilization might be expected to have greater explanatory power prior to this development. But this does not seem to be the case. We ran all three of the above vector autoregressions over the period 1950.I to 1971.II (we also excluded

the price controls period) and the explanatory power of capacity was even weaker. The null hypothesis that the coefficients of capacity are zero could not be rejected at a significance level below 33 percent.

If the aggregate supply curve is of the flat-steep type assumed here, we might also expect to find the output elasticity of price to be positively correlated with the rate of capacity utilization. One model consistent with this hypothesis is

$$(3) \quad \ln P_t = a + bT_t + \sum_{i=1}^8 c_i \ln P_{t-i} + \sum_{i=0}^8 d_i (C \ln X)_{t-i} + \epsilon_t$$

where

P = CPI . . . quarterly average.

X = real GNP.

This model provides no support for the view that capacity provides information about the slope of the supply schedule since the null hypothesis that the d_i 's are zero cannot be rejected at any reasonable level of significance (the F statistics is 0.71). This conclusion is subject to two important caveats. First of all the equation contains a contemporaneous value of $C \ln X$ on the right-hand side and therefore is subject to biased estimation. Secondly, there is an implicit assumption that the distribution of changes in aggregate demand are independent of the level of capacity utilization.

As a final test of the explanatory power of the rate of capacity utilization, we computed the residuals from the equation

$$(4) \quad p_t = a + bT_t + \sum_{i=1}^8 c_i p_{t-i}$$

We then regressed the squared residuals on C and found that C did not enter the regression in a significant way. The idea behind this test was that (4) represents the expected rate of inflation and that deviations

from expectations would be greater when prices move a lot (the supply curve is steep) than when they don't respond much due to a shift in demand. This test then is inconsistent with the view that the capacity utilization rate provides information about the slope of the supply curve.

Another way of testing the squared residuals of equation (4) is to define a dummy variable that takes the value 1 when the rate of capacity utilization is high and 0 when it is low. Then regress the squared residuals on this dummy. Using the value $C=85$ (a pure judgmental decision), we found that the dummy was a significant variable at the 5 percent level (but not at the 1 percent level).

Is the rate of capacity utilization a useful bit of information in explaining inflation? (In our judgment, the statistical evidence is, at best, ambiguous. The belief that stimulative aggregate demand policies may be pursued when the rate of capacity utilization is low since the danger of additional inflation is low requires more empirical support than we have found.

IV. Conclusion.

We have argued that the rate of capacity utilization is not a particularly useful piece of macroeconomic information. Conceptually its economic content is vague, and it cannot be used as a substitute for the concept of supply. It may be a compliment to supply, but the facts do not provide convincing support for this belief. Empirically, the case that the rate of capacity utilization is a useful bit of information in explaining inflation is ambiguous.

But even if the capacity utilization rate does provide information about the slope of the supply schedule, it is still of limited usefulness

as a guide to the conduct of aggregate demand policy. For if the supply schedule is shifting because of a policy action, its slope is only a second order of importance.