NBER WORKING PAPERS SERIES

PRODUCTIVITY, MARKET POWER AND CAPACITY UTILIZATION WHEN SPOT MARKETS ARE COMPLETE

Benjamin Eden

Zvi Griliches

Working Paper No. 3697

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1991

This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #3697
May 1991

PRODUCTIVITY, MARKET POWER AND CAPACITY UTILIZATION WHEN SPOT MARKETS ARE COMPLETE

ABSTRACT

Our test of price-taking behavior looks at the choice of capacity rather than the choice of output. It is motivated by a complete spot markets model in which goods are distinguished by the selling probabilities in addition to other characteristics. When output is explained by total man-hours and a capacity utilization proxy, the coefficient of the first variable is the elasticity of capacity with respect to fixed labor. Under competition and risk neutrality this coefficient is equal to an average labor share. We use this observation to interpret Abbot-Griliches-Hausman's regressions and to argue that once the capacity utilization proxy is included in the regression, Hall's data at the manufacturing level fail to reject the joint hypothesis of competition and risk neutrality. It is also argued that the coefficient of total man-hours does not tell us anything about monopoly power once the capacity utilization proxy is omitted from the regression.

Benjamin Eden
The University of Iowa
and
Technion-Israel Institute
of Technology

Zvi Griliches
Harvard University
and
NBER

Introduction

Recently Hall (1988, 1989) found that the short run elasticity of output with respect to labor is too high relative to the prediction of the standard competitive spot market model. Hall's empirical findings confirms previous findings by Okun and others. (See, Gordon for an excellent survey and additional references). Hall interprets his findings as evidence of monopoly power and/or increasing returns to scale. An alternative interpretation is that when output goes up both labor and capacity utilization go up. When we fail to measure the increase in capacity utilization, labor gets an "undeserved" credit for the increase in output and output elasticity with respect to labor appears to be "too high".

Abbott, Griliches and Hausman (A-G-H) examined the hypothesis that factor utilization is an omitted variable in the regression of the growth rate of output on the growth rate of labor input as measured by manhours per unit of capital services. They used hours per employee as a proxy for both capital and labor utilization rates and found that after the inclusion of this utilization variable in a simple OLS regression, the elasticity of output with respect to labor is not significantly different from the prediction of the standard competitive model.[1]

Hall assumes that the marginal product of capital is zero whenever capacity is not fully utilized and therefore labor should take all the credit for the increase in output. Here we defend the capacity utilization explanation against this objection and use a competitive model to interpret some of the key regressions in A-G-H.

The standard competitive spot market model defines goods by their physical characteristics and the place and time of delivery. The Arrow-Debreu (A-D) model defines goods by the state of nature in which delivery will occur. in addition to other characteristics. Here we define goods by the probability of sale which is the spot market analog of the state of nature characteristic in the A-D model.[2] We start with a version of Hall's first example (pages 928-929, in his JPE article).

We consider the problem of restaurants in a certain location who produce lunches. For simplicity we assume that fixed and variable labor are the only factors of production. Preparing a meal requires λ man-hours. Serving the meal and washing the dishes requires ϕ man-hours. The wage rate is constant and is given by w.

There are N_2 ex-ante identical potential buyers. Each potential buyer may wake up early or late. If he wakes up early he prepares lunch, take it in a brown bag and eat it at lunch-time. If he wakes up late he eats lunch at a restaurant. The number of agents who wake up late depends on the aggregate state which can not be observed: in aggregate state 1 N_1 buyers will wake up late and in aggregate state 2 all N_2 buyers will wake up late. Each buyer is willing to pay up to θ dollars for a meal if he did not bring lunch with him. It is assumed that θ is sufficiently large: $\theta > w\phi + w2\lambda$. The aggregate state is a serially independent random variable: in any given period, it may be 1 or 2 with equal probabilities of occurrence.

<u>The standard spot market model:</u> Goods are defined by physical characteristics and place of delivery only. There is a Walrasian auctioneer who uses a tatonnement type process to find the realization of demand. The auctioneer announces the price P_1 if the demand is low (state 1) and the price P_2 if the demand is high (state 2). There is a single price taking firm which chooses capacity, V, and output in state i, Q_i, to maximize expected profits:

$$(1)\ \max_V \ (\textstyle\sum_i \max Q_i(P_i - w\phi) \ ; \text{s.t. } Q_i \le V)/2 \ - w\lambda V .$$

The vector (P_1, P_2, Q_1, Q_2, V) is a competitive equilibrium in this framework if: (a) given the prices (P_1, P_2) the quantities

(Q_1, Q_2, V) maximize the expected profit (1) and (b) Q_i is equal to the number of buyers whose reservation price is above P_i.

We claim that equilibrium prices are given by:

(2) $P_1 = w\phi$; $P_2 = w\phi + w2\lambda$.

To see this claim note that for the prices (2), any $V \geq 0$, $Q_1 \leq V$ and $Q_2 = V$ is a solution to (1). Since the reservation price θ is sufficiently large, in equilibrium all customers will have a meal: $Q_1 = N_1$ and $Q_2 = N_2$.

In this standard competitive model there is excess capacity in the state of low demand and the price in this state is equal to SRMC = $w\phi$. In the state of high demand the price is higher than LRMC = $w(\phi + \lambda)$.

<u>The complete spot market model:</u> Goods are defined by the probability of sale in addition to other characteristics. There are two goods with the physical description "lunch". Let X_1 denote lunch that will be exchanged (sold) regardless of the aggregate state (with probability one) and X_2 denote lunch that will be exchanged if the aggregate state is 2 (with probability 1/2). There are two goods with the physical description "money": Y_1 denotes money that will be exchanged (received) regardless of the state and Y_2 denotes money that will be exchanged if the aggregate state is 2. The price of X_i in terms of Y_i is P_i. Risk neutrality implies that the shadow price of Y_2 in terms of Y_1 is 1/2. We shall use Y_1 as a numeraire.

Buyers arrive sequentially and can see all prices: they buy cheaper lunches first. We can think of the first group of N_1 buyers as participating in the market for X_1 and the second group of $N_2 - N_1$ buyers as participating in the market for X_2 if they do arrive. Thus the first market is opened with probability one and the second market is opened with probability 1/2.

Each firm (restaurant) chooses the number of lunches that will be sold regardless of the state, $x_1 \geq 0$, to maximize:

(3) $(P_1 - w\phi)x_1 - w\lambda x_1$;

and the number of lunches that will be sold only if the aggregate state is 2, $x_2 \geq 0$, to maximize the expected profit:

(4) $(P_2 - w\phi)x_2/2 - w\lambda x_2$.

The vector (P_1, P_2, x_1, x_2) is an equilibrium vector if: (a) given the price P_1 the quantity x_1 solves (3) and given the price P_2 the quantity x_2 solves (4) ; (b) demand is always satisfied in the markets which do open: x_1 is equal to the number of lunches demanded at the price P_1 in the first market and x_2 is equal to the number of lunches demanded at the price P_2 in the second market if it opens.

Can our equilibrium unravel? Sellers of commodity 2, after realizing a period of low demand may want to offer the good at an arbitrarily small price. In our example, this is too late: you can not sell another lunch to someone who already had lunch, even at a low price. In general, we assume that trade occurs sequentially and that buyers disappear when they complete their transactions, so that sellers find out that total demand is low only when it is too late to lower prices.

It is easy to see that equilibrium prices are:

(5) $P_1 = w\phi + w\lambda$; $P_2 = w\phi + 2w\lambda$.

At the prices (5) the firm is indifferent with respect to the choice of x_1 and the equilibrium quantities are demand determined.

Condition (5) says that the price P_1 is equal to LRMC expressed in terms of money that will be spent with certainty. That is in terms of Y_1 . Since risk neutrality implies that a dollar which is spent with certainty is equivalent to two dollars that are spent with probability half, the shadow price of Y_1 in terms of Y_2 is 2, and we can interpret (5) as saying that the price P_2 is equal to LRMC expressed in terms of Y_2 . To understand why only the LRMC is relevant let q_s denote the probability that market s will open (q_1 = 1, q_2 = 1/2) and let V_s denote the capacity allocated for the production of X_s. Then we can write the problems (3) and (4) as[3]:

$$(*)\ \max_V\ (\max_x\ (P_s - \phi w)x_s\ ;\ \text{s.t. } x_s \leq V_s\ ) - (w\lambda/q_s)V_s\ .$$

The solution to (*) must satisfy $x_s = V_s$ and therefore in terms of this enlarged commodity space, "capacity under-utilization" or "capacity over-utilization" do not occur: The cost of increasing production must include the cost of increasing capacity and therefore all prices are equal to LRMC.

We now express (5) in terms of the numeraire commodity[4] Y_1 :

$$(6)\ \ \text{ENR} = (P_1 - w\phi) = (P_2 - w\phi)/2 = w\lambda = \text{MCC}\ ,$$

where ENR denotes the expected net revenue per unit of capacity (net of operating cost) and MCC is the marginal capacity cost. Thus ENR is the same for both goods and is equal to MCC.

If we fail to distinguish between X_1 (lunches that will be sold with probability one) and X_2 (lunches that will be sold with probability less than one) we will erroneously reject competition because in the state of low demand the price is greater than

$SRMC = w\phi$ and in the state of high demand the average price is less than $w\phi + 2w\lambda$. Thus the complete spot market prices fluctuate less than the standard spot market prices.

Hall's test in the case of "capacity under-utilization": Hall argues that under competition and constant returns to scale the elasticity of output with respect to labor is roughly equal to the share of labor in total revenue. Let us assume that we observe the industry over two periods: in the first period N_1 lunches were sold and in the second period N_2 lunches were sold. Thus,

$$(7)\ dq = (N_2 - N_1)/N_1 \ ; \qquad dn = \phi(N_2 - N_1)/(\lambda N_2 + \phi N_1) \ ;$$

where dq is the percentage change in output, and dn is the percentage change in total man-hours. The observed elasticity is:

$$(8)\ dq/dn = (\lambda N_2 + \phi N_1)/\phi N_1 \ .$$

If the data were generated by the standard model then the price in the first period is $P_1 = \phi w$ and labor share in the first period is: $w(\lambda N_2 + \phi N_1)/w\phi N_1$ which is equal to the observed elasticity (8). If the data were generated by the complete spot market model, then the price of output in the first period is $w\phi + w\lambda$ and the labor share is therefore:

$$(9)\ \ w(\lambda N_2 + \phi N_1)/\ w(\phi + \lambda)\ N_1 \ ,$$

which is less than the observed elasticity (8).

Hall's markup is the observed elasticity (8) divided by the observed labor share. If the data were generated by the standard model Hall's markup is unity. If the data were generated by the complete spot market model Hall's markup is

(8)/(9) = $(\phi + \lambda)/\phi$, which is the ratio of long run marginal cost (LRMC) to short run marginal cost (SRMC). This ratio is always greater than unity and is large when the marginal capacity cost $w\lambda$ is large relative to the variable costs $w\phi$.

A numerical example may help. Assume that $w = 1$, $\lambda = 1/2$, $\phi = 1$, $N_1 = 6$ and $N_2 = 10$. In this case the price of the "first" 6 lunches is 1.5 and the price of the "last" 4 lunches is 2. The equilibrium data is described in Table 1.

period	output	labor input	average price	wage bill	revenue	share
1	6	11	1.5	11	9	1.22
2	10	15	1.7	15	17	.88

Table 1

Now dq = ln10 - ln6 = .51 and dn = ln15 - ln11 = .31. And dq/dn = 1.65 which is greater than (any average of) the share of labor in revenue .

The numerical example of Table 1 is repeated in Table 2, for the case in which prices are taken from the standard spot market model: 1 in the case of low demand and 2 in the case of high demand.

period	output	labor input	average price	wage bill	revenue	share	average share
1	6	11	1	11	6	1.83	1.29
2	10	15	2	15	20	.75	

Table 2

Note that prices in the complete spot market model fluctuate less than in the standard model. Since total revenue is the same in both models average price is also the same (26/16 = 1.625) [5,6]. Note also that in both Tables the ratio of the average wage bill to the average revenue is one but the average labor share is not (1.05 for Table 1 and 1.29 for Table 2).

Thus in the standard model there is one price in each period (the law of one price) and different prices across periods. In the complete spot markets model there is a difference in prices between the two "contingent" commodities within the same period, no difference in prices between periods, but no transactions in commodity 2 in the period of low demand. We get different prices for the "same" commodity within the same period and smaller differences in average prices across periods.

Which model is more "realistic" or more consistent with casual observations? The standard model assumes that trade is prohibited until all buyers arrive and total demand is known. This prohibition on trade is not only unrealistic but more importantly it is not in the spirit of the competitive paradigm. Moreover, whenever capacity is not fully utilized we do pay more than SRMC. Airlines are a good example: the price of a ticket paid by passengers in a half empty flight is much higher than the cost of the additional meal served on the flight.

Another way in which the two models may be compared is in the amount of "friction" assumed relative to a "frictionless" world. The standard model is often used as the bench-mark. This is rather strange since the prohibition on trade implied by the tatonnement process is itself a "friction". The Arrow-Debreu model seems a better bench-mark. It is shown in Eden (1990) that the transactions in the complete spot markets model can be interpreted as the execution of ex-ante Arrow-Debreu type contracts.

<u>More on Hall's test:</u> In Proposition 1 of his 1989 paper Hall shows that under the assumptions of constant returns to scale and competition, the Solow residual should not be correlated with variables that are uncorrelated with productivity changes, like military spending and oil price. Hall interprets his

empirical findings as showing that this invariance property "fails conspicuously" and concludes that "the assumptions Solow made in developing the now-standard approach to productivity measurement are clearly false." He then tries to find which of the underlying assumptions do not hold and suggests first monopolistic power and then increasing returns as well as other factors, as possible explanations for this phenomenon.

Differentiability is critical in deriving the invariance property (equation 2.2 in Hall's paper) but the lack of differentiability is not considered as a potential reason for its failure. To show that this may be the case, let us start with what Hall regards as a non-explanation. This is the case of overhead labor which is labeled as non-explanation 1 and is identical to our restaurants example.

Hall shows that the Solow residual is zero if we start from a period in which capacity is underutilized. In terms of our example, under the standard model (8)/(9) = 1, if we use the underutilization period as a base. This does not work if we take a period in which capacity is fully utilized as a base.

Using the data from Table 2, reveals that the logarithmic derivative, dq/dn = 1.65, does not depend on which period we take as a base.[7] Labor share in the period of low demand, 1.83, is roughly the same as dq/dn. But labor share in the period of high demand, .75 , is much lower than dq/dn. For the two periods taken together the wage bill equals total revenue and the average share so defined is unity, which is again lower than dq/dn.

The failure of the invariance property in this case is due to the fact that there are two separate processes in this example: the capacity generating process and the process of transforming capacity into output. Each process is well behaved but together they are analogous to the production function $Q = \min(F/\lambda\ , V/\phi)$, where here F is fixed or overhead labor and V is variable labor. This production function is not differentiable at the relevant points and therefore does not satisfy Solow's assumptions.[8]

To test for increasing returns to scale, Hall (1989) suggests a comparison between the elasticity and the share of labor in total cost: these should equal under constant returns to scale even in the presence of monopoly power. Since in our example the share of labor in total cost is 1 and the elasticity is greater than 1, this comparison will lead to the rejection of the constant returns to scale hypothesis, in spite of the fact that we have two constant returns to scale production processes. Thus, it seems that in the presence of "overhead labor" Hall's conclusions of significant market power and increasing retururns to scale can not be defended even under the standard model.[9]

MCC = ENR?

We generalize the above example to the case in which the number of buyers that will arrive is a discrete random variable $\tilde{N}$ which may take S possible realizations: $N_1 < N_2 < \dots < N_S$. Buyers arrive sequentially and do not hang around. That is once the first group of N_1 buyers finish trading it disappears and the second group of $N_2 - N_1$ may arrive and so on until all $\tilde{N}$ buyers finish to trade. All potential buyers have the same standard downward sloping demand schedule.

We define X_s as a good with physical characteristics X that will be exchanged (sold) if the realization of $\tilde{N}$ is greater than or equal to N_s. Similarly Y_s is a good with physical characteristics Y (money) that will be exchanged (received) when the realization of $\tilde{N}$ is greater than or equal to N_s. The probability that $\tilde{N}$ is greater than or equal to N_s is denoted by q_s.

Market s will open if $\tilde{N}$ is greater than or equal to N_s. That is, with probability q_s. If market s is opened then X_s will be exchanged with Y_s at the market clearing price P_s.

We study here the case in which each firm specializes in the production of a single good. This is a reasonable simplifying assumption because there are no benefits from producing all the goods together. The more "realistic" case in which each firm produces all goods with physical characteristics X is studied in the Appendix where we obtain the same main results as in the single good case.

The day (period) is divided into two. Capacity is generated in the first part. Demand may show up with probability q_s in the second part. It is assumed that $(V)^{1/\alpha}$ units of labor are required to create V units of capacity and $\phi(V)^{1/\alpha}$ units of labor are required to transform V units of capacity into V units of output. Employment must be chosen at the beginning of the day before demand is realized. A firm that wants to create V units of capacity will hire $L = (V)^{1/\alpha}$ units of labor input, where each unit of labor works one unit of time to create capacity and an additional ϕ units of time if demand is realized. The number of hours per unit of labor, H, is thus: 1 or $1 + \phi$.

A firm that produces and sell X_s at time t will solve:

$$(10)\ \max_L q_s[P_{st}(L_{st})^{\alpha} - W_t(1 + \phi)L_{st}] - (1 - q_s)W_tL_{st} ,$$

where W_t is the wage rate at time t. (For a more general labor contract, see the Appendix). The first order condition for (10) is:

$$(11)\ EW_t/ \alpha(L_{st})^{\alpha-1} = q_sP_{st} ,$$

where $EW_t = q_sW_t(1 + \phi) + (1 - q_s)W_t$ is the expected wage payment per employee. From the first order condition (11) we get:

$$(12)\ ETW/ETR = L_{st}EW_t/q_sP_{st}(L_{st})^{\alpha} = \alpha ,$$

where ETW denotes the expected total wage bill and ETR denotes the expected total revenue. ETW/ETR is approximately equal to the average labor share[10] and therefore (12) says that the elasticity of capacity with respect to fixed labor, α, is equal to an average labor share.[11]

Let ,

$$(13)\ g(H) = (H - 1)/\phi\ .$$

Since $g(1) = 0$ and $g(1 + \phi) = 1$, we can use $g(\ )$ as a measure of capacity utilization and write realized output Q_t as:

$$(14)\ Q_t = (L_t)^{\alpha} g(H)\ .$$

We shall use the approximation[12]:

$$(15)\ \ln\{g(H)\} = a + \beta \ln H\ ,$$

where $\beta > 0$ is the elasticity of capacity utilization with respect to hours per man. Under (15) we can write:

$$(16)\ dQ = \alpha dL + \beta dH\ ,$$

where $dZ = \ln Z_t - \ln Z_{t-1}$. Let n_t denote the total input of manhours. Since,

$$(17)\ dn = dL + dH\ ,$$

we can write:

$$(18)\ dQ = \alpha dn + (\beta - \alpha) dH\ .$$

Let S denote the average labor share. Then (18) and the first order condition

(12) imply that $\gamma = 1$ in the following equation:

(19) $dQ = \gamma S dn + (\beta - \alpha) dH$.

<u>The capacity mark-up coefficient:</u> To show that the parameter γ can be used to measure market power, consider the following monopoly problem:

(20) $\max_L \; q[P(L^{\alpha}) \cdot L^{\alpha} - W(1 + \phi)L] - (1 - q)WL$,

where subscripts were omitted and $P(L^{\alpha})$ denotes the inverse demand function in market s. The first order condition for (20) is:

(21) $EW/ \alpha L^{\alpha - 1} = qP(1 + 1/\eta)$,

where η denote the elasticity of demand. It follows that:

(22) $S = LEW/qPL^{\alpha} = \alpha\,(1 + 1/\eta)$,

and therefore:

(23) $\gamma = 1/\,(1 + 1/\eta)$.

To interpret γ, note that the left hand side of (21) is the full marginal capacity cost (FMCC): the cost of creating an additional unit of capacity and using it to make output whenever the market opens. Using ER = qP and (21) leads to:
γ = ER/FMCC = the ratio of the expected revenue per unit of capacity to the full marginal capacity cost, or the "capacity markup" for short.

<u>Demand driven economy:</u> OLS is appropriate when demand shocks dominate. We therefore start by assuming that changes in the probability distribution of $\tilde{N}$

cause changes in q_{st} and P_{st} and change in employment via the first order condition (11). Changes in the realization of $\tilde{N}$ causes changes in the number of hours per unit of labor. We measure dQ with an error e that is not correlated with dn and dH.

Abbot, Griliches and Hausman (A-G-H) used the (NIPA) data in Hall's paper for total manufacturing and started their discussion with the OLS regression:

$$(24)\ dQ = .0331 + \underset{(.10)}{1.57}\ Sdn.$$

To interpret (24) , note that it can be viewed as estimating (16) having assumed that $\alpha = \beta$. If however, $\alpha < \beta$ and (18) is right, then the coefficient of Sdn does not imply, necessarily, the presence of monopoly power. A-G-H proceeded by running the regression:

$$(25)\ dQ = .0241 + \underset{(.16)}{1.10}\ Sdn + \underset{(.66)}{1.74}\ dH.$$

Here the coefficient of Sdn is the capacity mark-up but this coefficient is not significantly different from unity. Thus the data can not reject the hypothesis that it comes from a competitive economy which is driven by demand shocks.

The implied estimate of the elasticity of capacity utilization with respect to hours per man, β , is roughly 2.5. To interpret this estimate let $h = H \cdot e$, where h are "true" hours per employee, H is measured hours and e is the fraction of measured hours actually worked, or effort . We implicitly assumed that H responds to demand shocks but e does not. It is more natural to assume that both H and e vary in response to demand shocks.

To illustrate, consider the more general case discussed in the Appendix, in which each restaurant produces all S goods. When 80% of capacity is utilized, waiters wait until all customers finish lunch but they are actually busy only 90% of the time. When demand is high and all capacity is utilized, waiters are busy all the time and when they finish serving the meals they help to wash dishes, staying at work 10% longer. Observed capacity utilization is up by 20%, hours per employee are up by 10% , effort is up by 10% and true hours are up by 20% . In this example, the elasticity of capacity utilization with respect to true hours is one and the elasticity of capacity utilization with respect to measured hours, or β, is 2. The elasticity of effort with respect to measured hours, η_{eH} , is 1.

More generally, if the elasticity of cpacity utilization with respect to true hours is one, then $\beta = 1 + \eta_{eH}$ and our estimate of $\beta = 2.5$ implies $\eta_{eH} = 1.5$. Thus a 1% increase in measured hours is associated with a 1.5% increase in effort.[13]

<u>Supply shocks:</u> We now assume that $(V/\theta)^{1/\alpha}$ units of labor are required for creating V units of total capacity, where θ is a supply shock realized at the time employment decision are made. If the market is opened then the firm will need an additional $\phi[V/\theta]^{1/\alpha}$ units of labor to transform capacity into output.

A firm that wants to create V units of capacity will hire $L = (V/\theta)^{1/\alpha}$ units of labor input, where each unit of labor works one unit of time to create capacity. The amount of additional time per unit of labor which is required to transform capacity into output is ϕ units of time per unit of labor.

The firm maximizes:

(26) $qP\theta L^{\alpha} - LEW$.

This leads to:

(27) $dQ = \gamma S dn + (\beta-\alpha)dH + d\theta$,

where the capacity mark-up coefficient is $\gamma = 1$. Now if the supply shock is an economy-wide shock, then it is reasonable to assume that changes in θ will not change the equilibrium allocation of fixed labor across industries (the expected wage per efficiency unit, EW/θ, remains constant and therefore changes in θ do not effect the solution to [26]). Since L is independent of θ it follows that both dn and dH are independent of $d\theta$ and therefore OLS is appropriate.

If θ is an industry specific shock and if labor is mobile across industries then dL is correlated with $d\theta$, leading to a correlation of both dn and dH with $d\theta$. In this case, if the demand shocks are correlated across industries then dGNP seems appropriate as an instrument: it is correlated with dn and dH and it is not correlated with $d\theta$. A-G-H ran the following regression using dGNP as an instrument:

$$(28)\quad dQ = .0269 + \underset{(.23)}{1.24\, Sdn} + \underset{(.66)}{1.38\, dH}$$

Therefore even under the interpretation that the productivity shock is industry specific we find that the capacity mark-up coefficient is not significantly different from 1.

<u>Risk aversion as a second line of defense:</u> We now introduce risk aversion and demonstrate that a capacity mark-up coefficient which is greater than one is consistent with the complete spot markets model: in the presence of risk aversion, the firm will require a risk premium on the installed capacity and therefore MCC < ENR.

To show this point let us assume that the firm uses the von-Neumann-Morgenstern utility function, U( ) to value profits and therefore solve:

$$(10') \max_L qU[PL^{\alpha} - W(1+\phi)L] + (1-q)U(-WL) ,$$

where U is strictly concave and differentiable. Let $U'(2) = U'[PL^{\alpha} - W(1+\phi)L]$ denote the marginal utility in the good state and let $U'(1) = U'(-WL)$ denote the marginal utility in the bad state. The first order condition for (11') is:

$$(11') \quad W'/\alpha L^{\alpha-1} = qP ;$$

where

$$(29) \; W' = qW(1+\phi) + (1-q)W(U'(1) / U'(2)),$$

is a "certainty equivalent" of the random wage payment. Note that (11') leads to:

$$(13') \; LW'/ETR = \alpha.$$

Let,

$$(30) \; \gamma = W'/EW$$

denote the ratio of the "certainty equivalent" wage payment to the expected wage payment. Concavity implies that $\gamma > 1$. Substituting (30) in (13') leads to:

$$(31) \; S = ETW/ETR = \alpha/\gamma .$$

Therefore the coefficient of Sdn in the regressions which includes both Sdn and dH as explanatory variables should be $\gamma > 1$.

The size of γ depends on the length of time for which employment is "fixed". If for example labor is fixed for a year then the risk premium on wage payment should be the risk premium on comparable stocks, say 7% which will lead to $\gamma = 1.07$.

Conclusions

Whenever restaurants do not operate at full capacity, prices are much higher than short run marginal cost. Hall's test interprets this type of observation as an indication of monopoly power. Here we assume that buyers arrive sequentially and explain the apparent discrepancy between marginal cost and price as the result of mis-specifying the commodity space: We define goods by the probability that they will be exchanged, as well as by other characteristics.

Our test of price-taking behavior looks at the choice of capacity rather than the choice of output. We define the capacity mark-up as the ratio of expected revenue per unit of capacity to the full marginal capacity cost, and test the joint hypothesis that the capacity mark-up is unity and agents are risk neutral.

When output is explained by total man-hours and a capacity utilization proxy, the coefficient of the first variable is the elasticity of capacity with respect to fixed labor and the capacity mark-up can be obtained by dividing this coefficient by an average labor share.

Abbot-Griliches-Hausman's OLS regressions are appropriate if the economy is driven by demand shocks: it convincingly shows that once the capacity utilization proxy is included in the regression, Hall's data at the manufacturing level fail to reject our joint hypothesis. OLS is also appropriate when the economy is driven by aggregate technological shocks to the capacity generating process, but it is not appropriate if the shocks are industry specific. Using the growth rate in GNP as an instrumental variable does not lead to a change in the main result.

Further research is required for generating a hypothesis about the size of the "capacity mark-up" coefficient in the presence of risk aversion. At the present we can say that this coefficient should be greater than unity even in the absence of monopoly power. A more powerful prediction requires the length of time for which labor is "fixed".

Another extension will allow for goods to differ both in the time of delivery and the probability of being sold: It will be interesting to study the behavior of a restaurant that can serve meals at any time during the day rather than only at lunch-time. Similarly, we should allow manufacturing firms to increase both the number of shifts and the number of hours per shift in response to a high realization of demand.

FOOTNOTES

1 They also used disaggregated plant-level data and energy inputs as a proxy for capacity utilization. When they disaggregate variable inputs into labor, materials and energy, they found that most of the "mark-up" is attached to the energy coefficient, supporting the above interpretation.

2 In this respect we follow Eden's work which develop a competitive version of Prescott (1975) and Butters (1977) model. Recently there has been a revival in the use of this type of models: Lucas (1989), Rotemberg and Summers (1990) and Woodford (1990). The paper by Rotemberg and Summers is especially relevant in the present context. It provides some indirect evidence which supports the approach used here.

3 Condition (*) expresses the profits in terms of Y_s. In particular, $w\lambda/q_s$ is the per unit capacity cost expressed in terms of Y_s. This is so because risk neutrality implies that the shadow price of Y_s in terms of Y_1 is q_s and λw is the cost in terms of Y_1.

4 Under risk neutrality, magnitudes expressed in terms of Y_1 are expected values. Thus for example the expected price of X_s is equal to the expected long run marginal cost for producing this good: $q_s P_s = q_s w\phi + w\lambda$. But unfortunately, the average price of all transactions can not be easily interpreted.

5 Note that average price is: $w\phi + w\lambda + w\lambda(N_2-N_1)/(N_1 + N_2)$. Therefore if we use $(N_2-N_1)/(N_1 + N_2)$ as a measure of demand variability, we may say that average price (and average cost per unit sold) is lower when demand is smoother.

[6] When the individual demand function is downward sloping, average price in the complete spot markets model is higher relative to the standard model. To see this point, note that the quantity sold to the first N_1 buyers will be smaller under the complete spot markets model and therefore the measure of demand variability (in footnote 5) is larger.

[7] The arc elasticities will differ depending on where they are evaluated at. With 1 as a base , it is 1.83; with 2 it is 1.5. At the ratio of the averages it is 1.63. At the average of the ratios it is 1.67, close enough to the logarithmic derivative.

[8] It seems that non-differentiability is required to maintain Hall's assumption of zero shadow price of capital in periods of underutilization.

[9] Hall admits other explanations. These are of two types: other market imperfections and measurement errors. External effects, monopsony power and changes in work effort which do not change workers' compensations falls in the first category. The last item falls in this category because perfect competition requires perfect information and in this case workers' compensation must be a function of effort.

[10] Because of Jensen's inequality, our definition of average labor share ETW/ETR is lower than the average of wage payments divided by revenues E[TW/TR] which is the definition of s used by A-G-H. In practice we expect that the difference between the two definitions is not large.

[11] To illustrate (12) note that in the numerical example of Table 1, the elasticity of output with respect to fixed labor is one and the ratio of the average wage bill to the average revenue is 13/13 = 1.

[12] $\ln[g(1)]$ is not defined. This is not a problem when the firm produces all goods and total output is always positive. See the Appendix.

[13] Fay and Medoff (1985) conducted a survey which may be used to learn about the size of η_{eH}. They found that going from the deepest trough to "normal" times is associated with a 31% increase in production, 23% increase in total blue colour man-hours and 8% increase in effort, defined as the fraction of time devoted to regular production. Okun's law suggests that a 31% increase in output is associated with a 12% increase in employment, so that hours per employee were up by only 11%. This suggests an elasticity of effort with respect to measured hours of 8/11, which is smaller than what we get.

Fay and Medoff asked about large changes: going from the deepest trough to "normal" times . The elasticity for small changes is likely to be higher than 8/11, becuse the firm may adjust measured hours to large changes but not necessarily to small changes: it may let workers spend less effort when demand is mildly low but close a shift and demand higher effort when the drop in demand is large. In addition, labor hoarding may be more pronounced among white colour workers. These considerations suggests that an elasticity of 1.5 is not unreasonable.

REFERENCES

Abbott, T. A. III, Griliches, Z., and Hausman, J.A., "Short Run Movements in Productivity: Market Power versus Capacity Utilization", presented at the NBER Summer Institute (Aug. 1988).

Butters, G., "Equilibrium Distributions of Sales and Advertising Prices," Review of Economic Studies 44: 465-491 (1977).

Eden, B. "Marginal Cost Pricing When Spot Markets are Complete," Journal of Political Economy, Dec. 1990, Vol. 98, No. 6, 1293-1306.

________"Trading Uncertainty, Markups and Measured Productivity: A Comment on Hall's paper." Carnegie-Rochester Conference Series on Public Policy, Edited by: K. Brunner and A. Meltzer, Vol. 27 (Amsterdam , North Holland), 1987.

Fay J.A. and Medoff J.L., "Labor and Output Over the Business Cycle: Some Direct Evidence" American Economic Review , Vol. 75, No. 4 (Sep. 1985) ,638-655.

Gordon Robert,J. "Are Procyclical Productivity Fluctuations a Figment of Measurement Error" mimeo, October 1990.

Hall, R. E., " The Relation between Price and Marginal Cost in U.S. industry," The Journal of Political Economy, October 1988, Vol. 96, No. 5, 921- 947.

________ " Invariance Properties of Solow's Productivity Residual" NBER working paper #3034, July 1989.

Lucas, R. E., Jr. "The Effects of Monetary Shocks When Prices are Set in Advance," mimeo, University of Chicago, November 1989.

Okun, Arthur M. "The Gap between Actual and Potential Output", Proceedings of the American Statistical Association, 1962. Reprinted in Edmund S. Phelps, ed., Problems of the Modern Economy. New York: Norton, 1965.

Prescott, E. C., "Efficiency of the Natural Rate," Journal of Political Economy, 83 (Dec. 1975): 1229-1236.

Rotemberg, J.J., and L.H. Summers, "Inflexible Prices and Procyclical Productivity," The Quartely Journal of Economics , Nov. 1990.

Woodford, M., "Optimal Monetary Policy in an Economy with Sequential Service of Buyers," mimeo, University of Chicago, July 1990.

Appendix : The problem of a firm that produces all goods with physical characteristics X

As before $V^{1/\alpha}$ units of labor are required for creating V units of total capacity. But here total capacity is allocated among the S markets. Let μ_s denote the fraction of total capacity allocated to market s. Thus $V_s = \mu_s V$ is the amount of capacity allocated to market s. If j markets are opened then the firm will need an additional $\phi(\sum_{s\leq j} \mu_s)V^{1/\alpha}$ units of labor to transform $V\sum_{s\leq j} \mu_s$ units of capacity into output. Thus, there is a constant returns to scale in the process of transforming capacity into output, but there is a decreasing returns to scale in the production of capacity.

Employment must be chosen at the beginning of the day before demand is realized. Therefore a firm that wants to create V units of capacity will hire $L = (V)^{1/\alpha}$ units of labor input, where each unit of labor works one unit of time to create capacity. The amount of additional time per unit of labor which is required to transform capacity into output is $\phi\sum_{s\leq j} \mu_s$ units of time per unit of labor if j markets are opened.

It is assumed that the firm offers a labor contract that specifies total compensation and the amount of work per unit of labor as a function of the demand shock:

(A5) $W(\tilde{N}) \;\; ; \phi(\tilde{N})$.

Let, $EW = \sum_s \text{Prob}(\tilde{N} = N_s)W(N_s)$, denote the expected wage per unit of labor. The firm's problem is to choose V_{st} and L_t to maximize:

(A6) $\sum_s q_s P_{st} V_{st} - L_t EW \;\; ; \quad \text{s.t.} \;\; \sum_s V_{st} = (L_t)^\alpha \; ;$

where q_s = Prob($\tilde{N} \geq N$). Substituting the constraint in the objective function we get:

(A7) $(L_t)^{\alpha}\sum_s q_s P_{st}\mu_{st} - L_t EW$.

The first order condition for (A7) are given by :

(A8) $\alpha L^{\alpha-1} ER = EW$;

where $ER = \sum_s q_s P_{st}\mu_{st}$ is the expected revenue per unit of capacity . Multiplying both sides of (A8) by L and using $V = L^{\alpha}$ leads to:

(A9) ETW/ ETR = α,

Thus in this model, the average labor share is equal to the elasticity of capacity with respect to employment, α. Here,

(A10) $g(H) = (H - 1) / \phi = \sum_{s \leq j} \mu_s$.

We can therefore write $Q_t = (L_t)^{\alpha} g(H)$ as in the single good case and follow the same procedure as before.