Capacity Utilization under Increasing Returns to Scale

Yi Wen
Department of Economics
Hong Kong U. of Science & Technology
Clear Water Bay, Kowloon
Hong Kong
\[
\max E \sum_{t=0}^{\infty} \beta^t \left( \log(c_t) - \frac{n_t^l + \gamma}{1 + \gamma} \right)
\]

s.t. \( c_t + k_t + 1 - (1 - \delta_t) k_t = A_t \bar{e}_t(u_t k_t)^{\alpha} n_t^{1 - \alpha} \);  

\[
\delta_t = \tau u_t^\theta, \quad 0 < \tau < 1, \theta > 1;
\]

\[
\bar{e}_t = (\bar{u}_t k_t)^{\alpha \eta} n_t^{(1 - \alpha) \eta}
\]

where \( \gamma \in (0, 1), \eta \in (0, 1) \); and \( A_t \) is technology, \( u_t \in (0, 1) \) is the rate of capacity utilization, \( \delta_t \in (0, 1) \) is capital depreciation defined as an increasing function of capacity utilization, and \( \bar{e} \) is the productive externality as a function of the average economy-wide levels of productive capacity and labor. The restriction \( \theta > 1 \) is meant to impose a convex cost structure on capacity utilization and assure interiority.

The first order condition with respect to capacity usage \( u_t \) is

\[
\alpha \frac{\gamma}{u_t} = u_t^{\theta - 1} k_t
\]

where the LHS is the marginal output (\( y \)) gained by increasing the capacity utilization rate (\( u \)), and the RHS is the marginal loss in terms of
capital depreciation due to the intensified usage of existing capital stock. Equation (2) can be rewritten as:

\[ ut = \left( \frac{\alpha y_t}{k_t} \right)^{\frac{1}{\theta}}, \]  

which says that the optimal rate of capacity utilization is determined by the marginal product of capital. In other words, capital should be used more intensively during economic booms when its marginal product is high and less intensively during recessions when its marginal product is low. Using equilibrium conditions, one can use (3) to obtain an expression for the optimal capacity utilization rate in terms of aggregate capital and labor:

\[ ut = \frac{1}{\alpha A t k_t^{\alpha(1+\eta)} - n_t^{1-\alpha}(1+\eta) \theta - \alpha(1+\eta)}. \]

Notice that capacity utilization is homogenous with degree zero in capital and labor only if the externality is zero. Otherwise, it is homogenous with a degree greater than zero.
Finally, substituting (4) into the production function, we have

\[ y_t = b \ A_t^\tau n \ k_t^\alpha(1+\eta)\tau k \ n_t^{(1-\alpha)(1+\eta)\tau n} \]  

(5)

where the constants \( b, \tau_k, \tau_n \) are defined as (assuming \( \theta - \alpha(1+\eta) > 0 \)):

\[ b = \alpha \left( \frac{\alpha(1+\eta)}{\theta - \alpha(1+\eta)} \right), \quad \tau k = \frac{\theta - 1}{\theta - \alpha(1+\eta)}, \quad \tau n = \frac{\theta}{\theta - \alpha(1+\eta)}. \]  

(6)

Expression (5) is the reduced-form aggregate production function, which indicates that capacity utilization effectively alters the equilibrium production function and amplifies technology shocks. These changes are mainly reflected by \( \tau_k \) and \( \tau_n \). When the externality \( \eta = 0 \), it is easy to show that \( \tau_k < 1 \) and \( \tau_n > 1 \) since \( \alpha < 1 \) and \( \theta > 1 \).
Two effects are worth stressing. First, when the economy exhibits constant returns to scale ($\eta = 0$), capacity utilization has no effect on aggregate returns-to-scale (i.e., the factor elasticities sum to one):

$$\alpha \tau_k + (1 - \alpha) \tau_n = 1;$$

(7)

but it has significant effects on the distribution of factor elasticities: the capital elasticity decreases and the labor elasticity increases (because $\tau_k < 1$ and $\tau_n > 1$). For example, suppose the capital share of national income $\alpha = 0.3$ and the depreciation elasticity parameter $\theta = 1.4$ (which is the value calibrated by Greenwood et al. [31] according to the steady-state rate of capital depreciation $\delta = 0.025$ and the time discount factor $\beta = 0.99$), these then imply $\tau_k = 0.36$ and $\tau_n = 1.27$, which means that the effective capital-output elasticity is just about 0.1 while the actual capital-output elasticity is 0.3, and the effective labor-output elasticity is around 0.9 while its actual value is 0.7. This provides a possible explanation for the apparent empirical puzzle that the estimated capital elasticity is near zero and the estimated labor elasticity is near
This `elasticity effect' of capacity utilization is consistent with recent empirical findings of Shapiro [46] and Burnside, Eichenbaum and Rebelo [21]. It arises because capacity utilization tends to co-move with labor and counter-move with capital (see equation 3). The reason is that the net marginal gain of capacity utilization is an increasing function of labor but a decreasing function of the capital stock at the steady state (due to the fact that capacity utilization accelerates the depreciation of existing capital stock). Thus, in addition to the direct multiplier effect of capacity utilization on amplifying technology shocks (from equation 5 one can see that this effect is $\tau_n$), there is also an indirect multiplier effect resulting from the positive `elasticity effect' of capacity utilization on labor, which further amplifies technology shocks as it effectively increases the responsiveness of the production level to these shocks (remember that capital stock is fixed in the short term, so the adverse `elasticity effect' on capital does not matter).
Secondly, when the economy is subject to mild increasing returns to scale ($\eta > 0$), capacity utilization not only alters further the equilibrium distribution of factor elasticities, but also has an effect on the aggregate returns-to-scale, since, if $\theta - \alpha(1+\eta) > 0$,

$$\alpha(1+\eta)\tau_k + (1-\alpha)(1+\eta)\tau_n > (1+\eta).$$

(8)

This is called the `returns-to-scale effect' of capacity utilization. It can be shown that the `elasticity effect' of capacity utilization alone is not sufficient for explaining the fact that the estimated labor-output elasticity often exceeds one. The `elasticity effect' and the `returns-to-scale effect' combined together, however, are able to explain this well-known empirical puzzle. For example, given the previous parameterization, a mild degree of externalities in production, $\eta = 0.11$, would result in an observed labor-output elasticity around 1.02 while the actual elasticity is less than 0.78.
3. Solving the Model

These first order conditions are:

\[ ct = (1 - \alpha) \frac{y_t}{n_t^{(1+\gamma)}}. \]  
(9)

\[ \delta t = \frac{\alpha y_t}{\theta k_t}. \]  
(10)

\[ l = \beta E_t \frac{ct}{ct+1} \left( \alpha \frac{y_{t+1}+1}{k_{t+1}} + (1 - \delta_{t+1}) \right). \]  
(11)

\[ ct + kt+1 - (1 - \delta_t) k_t = y_t = b A^\tau n_k^{(1+\eta)^\tau k} n_t^{(1+\eta)^\tau n}. \]  
(12)

plus a transversality condition. The first equation determines the labor market equilibrium, the second equation determines the optimal rate of capacity utilization (or capital depreciation), the third one is the consumption Euler equation, and the last one is the consumer's budget constraint expressed at the optimal rate of capacity utilization using the reduced-form production function (5). From equations (10) and (11), we derive the relationship in the steady state between the depreciation elasticity parameter \( \theta \) and other structural parameters in the model:
\[
\theta = \frac{(1 - \beta(1 - \delta))}{\beta \delta}
\]  

(13)

where \(\delta\) is the steady-state rate of capital depreciation. It is easy to verify that \(\theta > 1\) if \(\delta \in [0,1]\) and \(\beta < 1\).

Equation system (9)-(12) does not have analytical solutions. Instead, we characterize the model's dynamics by linearizing these first order conditions around the steady state following King, Plosser and Rebelo [35]. Using hat variables to denote the linearized variables, one can show that the system (9)-(12) can be reduced to the following linear dynamic systems under rational expectations:

\[
\begin{bmatrix}
\hat{k}_{t+1} \\
E(\hat{c}_{t+1}|t) \\
\end{bmatrix}
= B 
\begin{bmatrix}
\hat{k}_t \\
\hat{c}_t \\
\end{bmatrix}
\]

(14)

\[
\begin{bmatrix}
\hat{y}_t \\
\hat{i}_t \\
\hat{u}_t \\
\end{bmatrix}
= \Pi 
\begin{bmatrix}
\hat{k}_t \\
\hat{c}_t \\
\end{bmatrix}
\]

(15)

where \(B\) and \(\Pi\) are real matrices, and technology \(A_t\) has been assumed to be constant for the purpose of exposition.
4. Necessary and Sufficient Conditions for Indeterminacy

\[-1 < \det(B) < 1\]
\[-(1 + \det(B)) < \text{tr}(B) < (1 + \det(B)).\]  \hspace{1cm} (17)

\[
\det(B) = \frac{1}{\beta} \left( 1 + \frac{\eta(1+\gamma)(1-\beta)\tau_n}{(1+\gamma) - \beta(1-\alpha)(1+\eta)\tau_n} \right), \hspace{1cm} (18)
\]

\[
\text{tr}(B) = 1 + \det(B) + \frac{(1+\gamma)(1-\beta)(\theta\alpha)(1-\alpha(1+\eta)\tau_k)\delta}{(1+\gamma)\beta(1-\alpha)(1+\eta)\tau_n}, \hspace{1cm} (19)
\]

where $\tau_k$ and $\tau_n$ are defined in (6). Notice that when an externality does not exist ($\eta = 0$), the determinant is simply $1/\beta (> 1)$, indicating a saddle-path-stability as in a standard RBC model.

Notice that $\tau_k$ and $\tau_n$ are positive and finite when $\eta = 0$. To maintain this property in the presence of externalities, we restrict our analyses to cases where

\[
\alpha(1+\eta) < \theta. \hspace{1cm} (20)
\]

The common denominator in expressions (18) and (19) suggests that
when the externality parameter $\eta$ increases from zero, the model may go through a point of discontinuity at which $\text{det}(B)$ and $\text{tr}(B)$ both change sign, passing from $+\infty$ to $-\infty$, if the conditions $\tau_n > 0$ and $\alpha(1+\eta)\tau_k < 1$ still hold.

To reach the discontinuity point, however, we also need

$$\beta(1-\alpha)(1+\eta)\tau_n \geq (1+\gamma).$$

(22)

This is an important necessary condition for indeterminacy. It is analogous to that derived by Benhabib and Farmer [11] in a continuous time model without variable capacity utilization, since the left-hand side of it is nothing but the effective labor-output elasticity evaluated at the optimal capacity utilization rate. This condition hence has a simple interpretation: the equilibrium wage-hours locus in the labor market need to be positively sloped and to cut the labor supply curve from below to generate indeterminacy. This condition implies:

$$\eta > \frac{\theta(1+\gamma - \beta(1-\alpha)) - (1+\gamma)\alpha}{\beta(1-\alpha)\theta + (1+\gamma)\alpha}.$$  

(23)

It needs to be verified that requirement (23) and requirement (21) are compatible for certain parameter values of the model imply:
\[ \beta(1-\alpha)(1-\beta(1-\delta)) > (1+\gamma)(1-\beta)\alpha . \]  

(24)

It is obvious that there exist regions of the parameter space such that (24) is satisfied (e.g., for \(\beta\) close to one, or for \(\alpha\) small enough).

Necessary and sufficient conditions for indeterminacy are (27), (26) and (21). To show that the parameter region thus specified for \(\eta\) is not empty for realistic parameter values, consider the parameterization of Farmer and Guo [28] in a quarterly model: \(\gamma = 0\) (Hansen's [33] indivisible labor), \(\alpha = 0.3\), \(\beta = 0.99\) and \(\delta = 0.025\) (implying \(\theta = 1.4\)). Figure 1 shows regions of indeterminacy as functions of \(\eta\). It is seen that the permissible regions for indeterminacy are very large.

From the necessary condition (22), it is evident that this model requires a smaller degree of increasing returns than the model with fixed capacity utilization (Benhabib-Farmer [11]) to induce indeterminacy due to the presence of the term \(\tau_n\), which arises because of the effect of capacity utilization. At the above parametrization, the required value of \(\eta\) for inducing indeterminacy is 0.1036, implying a downward sloping
aggregate labor demand curve (slope = (1-α)(1+η)-1 = -0.23) and a mild increasing return-to-scale (1+η = 1.1). If capacity utilization were fixed, the minimum degree of externality for generating indeterminacy would be 0.4927, implying a substantially larger return-to-scale (1+η = 1.5) and an upward sloping aggregate labor demand curve.

The insight is that variable capacity utilization increases the elasticity of output with respect to labor and that, in the presence of mild external increasing returns to scale, this effect can be sufficient to push the labor elasticity of output above one. In equilibrium the marginal product of labor is thus increasing in labor rather than decreasing as in the standard model. And as shown by Benhabib and Farmer [11], this implies that the rational-expectations equilibrium is indeterminate.
The intuition can be understood using the labor supply and demand curves (Aiyagari [2]). Since the capacity utilization rate can respond to changes in consumption level at the impact period, an upward shift of the labor supply curve caused by an increase in the initial consumption level can at the same time trigger an upward shift of the aggregate labor demand curve (because of increases in the capacity utilization rate). If this shift in the labor demand curve is large enough, then equilibrium labor and real wages will both increase (indicating a rise in the permanent income), which substantiates the initial increase in consumption. If capacity utilization were fixed, however, an upward shift of the labor supply curve would result in a decrease in equilibrium labor and output unless the aggregate labor demand curve was upward
sloping and was steeper than the labor supply curve (as in the model of Benhabib and Farmer [11]. Thus, capacity utilization explains why multiple equilibria may emerge in a one-sector growth model with a downward sloping aggregate labor demand curve.
Table 1. Sample and Population Moments

<table>
<thead>
<tr>
<th>Var</th>
<th>$\sigma_x/\sigma_y$</th>
<th>Cor(x, y)</th>
<th>Autoc</th>
<th>$\sigma_x/\sigma_y$</th>
<th>Cor(x, y)</th>
<th>Autoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>1.00</td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>0.74</td>
<td>1.00</td>
<td>0.95</td>
<td>0.04</td>
<td>1.00</td>
<td>0.79</td>
</tr>
<tr>
<td>inv</td>
<td>2.63</td>
<td>0.88</td>
<td>0.97</td>
<td>4.63</td>
<td>0.37</td>
<td>0.93</td>
</tr>
<tr>
<td>n</td>
<td>0.88</td>
<td>0.87</td>
<td>0.92</td>
<td>0.99</td>
<td>0.99</td>
<td>0.79</td>
</tr>
<tr>
<td>u</td>
<td>-</td>
<td>0.62</td>
<td>0.92</td>
<td>0.76</td>
<td>0.99</td>
<td>0.79</td>
</tr>
<tr>
<td>y/n</td>
<td>0.83</td>
<td>-</td>
<td>-</td>
<td>0.04</td>
<td>0.96</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Note: Variables (y, c, inv, n, u, y/n) stand for output, consumption, investment, labor, capacity utilization, and productivity respectively. The U.S. data were pre-detrended by a quadratic time trend.