*Department of Economics*
*Clemson University*

# THE STRUCTURE OF COSTS AND THE COMPETITIVE EQUILIBRIUM

There are three cases that we wish to study. Each involves the equilibrium adjustment in a competitive market to changes in exogenous constraints. Number one is the effect of a change in an input price; two is the effect of an input restriction; and three is the effect of an output quota.

An example of the first case is the minimum wage. When the government increases the minimum wage, we want to be able to describe its effect on the Average Cost functions of the firms using this resource, unskilled labor. If the minimum wage affects AC, it must also have an effect on the price of output. An increase in the minimum wage must increase AC and hence increase the price of output. Thus, the level of market output must fall. The question becomes one of how the optimal firm size changes and how the optimal number of firms adjusts. That is, do firms enter or leave the market? A collateral question is, How is the use of other inputs affected by the change in the price of the one? We will attempt to answer all of these questions as we develop the theory.

But, you have to crawl before you walk, as they say. So, we will first take up the simplest question first: the market equilibrium effect of output quotas.

## Quotas

Quotas are common enough as a market phenomenon. Government often imposes such restrictions as, for instance, in agriculture. In the U.S., both peanuts and cotton have or have had such restrictions.

Consider Figure 1. Shown there is a standard analysis of competitive equilibrium depicted by a market demand curve and the long-run average and short-run cost curves for one of the many identical firms. The initial equilibrium is $\{P_0, Q_0\}$ which is an equilibrium based on the assumption that the number of firms, $n$, is such that $nq_0 = Q_0$. Thus, all firms produce where price is equal to their marginal cost; they are behaving according to the objective of maximizing profit. Moreover, market price is equal to minimum average cost so there are no excess profits; thus, there is no incentive for firms to enter or leave the market.

Next consider the effect of a government quota on output. Suppose government decrees that market output must be cut from $Q_0$ to $Q_1$. Clearly some enforcement mechanism must accompany this decree because in the absence of such, the quota will create a disequilibrium. If market output is reduced and market price increases, firms will have an incentive to produce more, not less.

Assume that government allocates the quota level of output uniformly across the $n$ firms in the industry. That is, each firm gets the right to produce $[Q_1 / Q_0]$ of the amount that it produced before the quota. Thus, each firm has the right to produce $q_1$ units of output as shown in Figure 1.

The firms are now not maximizing profits. However, they are making excess profits. In the short run, profits are the distance $ab$ times $q_1$. If this situation prevails into the long run, the firms have an incentive to change their production technology and move to a smaller plant size, one where short run average cost is equal to $c$. This increases profits from $ab$ times $q_1$ to $ac$ times $q_1$, again assuming that the quota system prevails into the long run.

Finally, let's consider what happens if the quotas that are initially allocated on a pro rata basis become transferable. In such an event, there are gains from trade. One quota unit (i.e., one unit of output) is worth only *ab* on average. However, at the margin it is worth *ab'*. Hence, some quota holders will be induced to sell their production rights and leave the market. This process will continue until all the firms that remain in production return to the production level $q_0$. At this point, the quota rights sell for ($P_1$-$P_0$) per unit. Firms have no incentive to acquire additional units. For instance the value of additional units past $q_0$ is only *de* as shown in Figure 1, which is less than their price.

It is important to recognize that while the firms that remain in production started out with an initial share of quota units for which they paid nothing, this has no effect on their decision making. Each firm has the opportunity to sell its rights. Hence, from wherever they came, they must be valued at the market price for quota units. Profits from production is calculated on the basis of the market prices of inputs including the market price of quota units. Hence, at the new equilibrium of {$P_1$,$Q_1$}, excess profits are zero. All of the profits that were created by the government restriction on output are impounded into the price of quotas. The firms that remain in production enjoy rents from their initial distribution of quotas in the amount of ($P_1$-$P_0$) times $q_1$, while the amount ($P_1$-$P_0$) times ($q_0$-$q_1$) is a payment for the quotas that are owned by the firms that cease production.

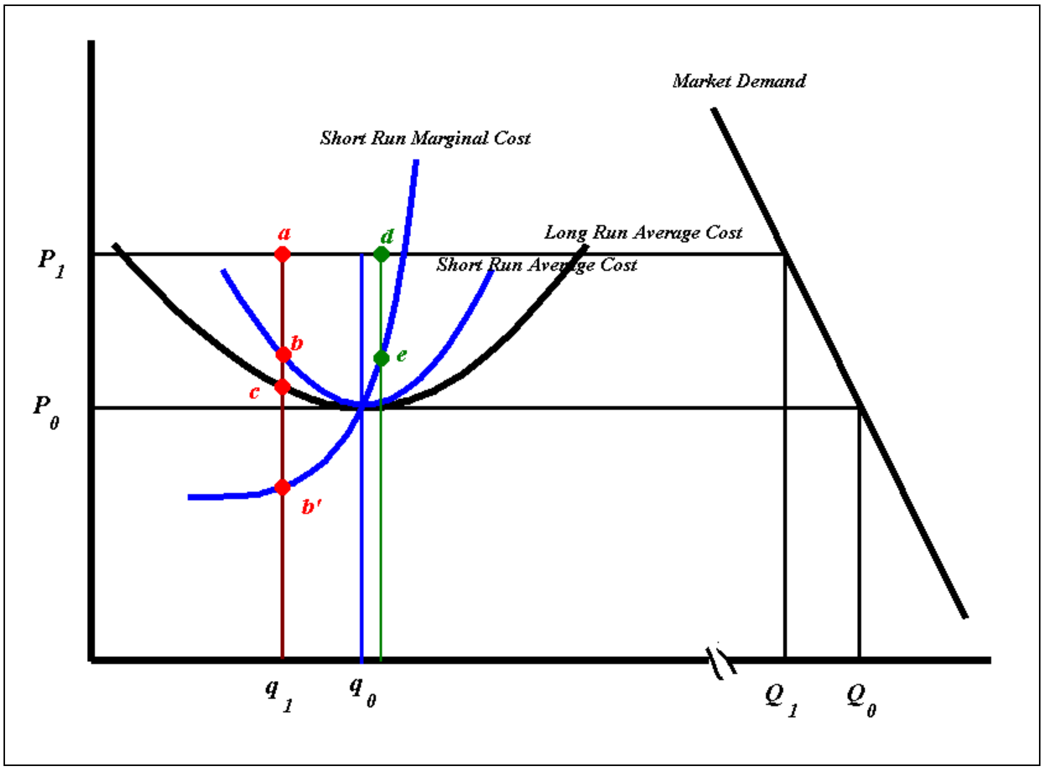


Figure 1

## Input Restrictions and Average Cost[1]

The next step in the process is to consider the effect of an input restriction. An application of this can be found in the study of regulatory policy. Buchanan and Tullock made an

[1] Maloney, M.T. and McCormick, R.E. "A Positive Theory of Environmental Quality Regulation" in Journal of Law and Economics vol 25 1982. p 99-123

insightful comment concerning environmental quality regulation. They noted that there is a decided disinclination toward pollution taxes in the practice of policy. Their explanation for this is that pollution restrictions are also output restrictions. Output restrictions raise price. While taxes can achieve this and transfer the pollution rents to the government, pollution abatement regulations can achieve the same end and transfer the rents to the polluting firms. This is the same conclusion that we developed above in the context of quotas.

B&T assumed a fixed proportions production function so that a restriction on an input was identical to a quota on output. However, in general the production function is not fixed proportions and an input restriction is arguably different from a quota. The purpose of this discussion is to examine the case of an input restriction in the more general case. The full discussion with some empirical evidence is found in Maloney and McCormick, "Environmental Quality Regulation," *JLE*, '82.

Assume that an industry is competitive with identical firms and a flat long-run industry supply curve. Assume that the firms produce with U-shaped average cost. Assume, further, that one of the $n$ inputs that make up average cost is pollution. Call it the $j$th input. Let it have zero price. In the pre-regulation world, firms pollute to the point where the marginal product of pollution is zero. The pre-regulation demand for pollution as an input is

$$x_j^*(w_1,\ldots,w_j,\ldots w_n,q)$$

where the normal notation applies, specifically, $q$ is firm level output.

In the B&T framework, the regulatory authority has the option of imposing a pollution tax, that is, charging a positive price $w_j > 0$, or imposing a pollution restriction, $\bar{x}_j$, where

$$\bar{x}_j < x_j^*(\{w_i; w_j = 0\}, q^*).$$

In other words, the pollution restriction is less than the firm is currently using. The current use of pollution is the firm's unrestricted demand for pollution evaluated at the competitive equilibrium output level, which is the quantity of minimum average cost denoted by $q^*$.[2]

We know that these two options have different effects on average cost. The pollution tax shifts the average cost function of the firm. It increases average cost and can shift the minimum of average cost either to the right or to the left depending on the elasticity of the pollution input demand with respect to the firm's output level.

The pollution restriction on the other hand shifts the firm's average cost, but under fairly ordinary assumptions causes it to bounce along the envelope of the unrestricted average cost. The effect is just like any other input restriction. The picture of a pollution restriction is nearly the same as the picture of the set of plant-size choices facing the firm along its long-run average cost planning curve.

Formally, if the output effect for the demand for input $j$ is positive, i.e.,

$$\frac{\partial x_j^*}{\partial q} > 0$$

[2] The price of pollution does not have to be zero prior to the regulation that restricts its use, but it emphasizes the point.

and continuous, then the pollution restricted average cost function touches the unrestricted average cost curve at some point. Since the pollution control regulation forces the firm to use less pollution than it would otherwise employ, the shift in average cost is to the left. This means that the minimum of the restricted average cost curve lies to the left of the minimum of unrestricted average cost. The shift in restricted average cost starts at the minimum of unrestricted average cost and the slope of the change in minimum average cost with respect to an input restriction starts at zero. In other words, if you draw the locus of minimum average cost with respect to changes in the input restriction, it, too, is U-shaped.

Consider, then, what happens to the market equilibrium and profitability of the firms in the industry as the regulatory authorities impose pollution control. Assume that the regulatory authorities mandate pollution control at the firm level and restrict the entry of new firms. The possibility theorem results:

> *For downward sloping market demand, there is some pollution restriction that will cause the profits of the existing firms to become positive.*

The rule is that there is some level of pollution control that will make the existing firms better off, though too much may make them worse off.

The market adjustment process is identified by the follow system of equations

$$P = D(nq)$$
$$P = MC(q, \alpha)$$
$$\Pi = [P - AC(q, \alpha)]nq$$

The first equation is the market demand clearing equation. The second is the firm-level profit maximization reaction function. The third is the profit level for the industry. Initially, market profit are zero by assumption of a pre-regulation competitive equilibrium.

The system is differentiated w.r.t. $\alpha$. This is the shift in marginal and average cost that is occasioned by the input restriction. The system is solved for $\frac{\partial \Pi}{\partial \alpha}$. This yields:

$$\frac{\partial \Pi}{\partial \alpha} = nq\left[ MC_\alpha \left( \frac{nD'}{nD' - MC_q} \right) - AC_\alpha \right]$$

For this to be positive, that is, for industry profit to increase as a result of the cost restriction, it is necessary that

$$\frac{1}{MC_q}\left[ \frac{MC_\alpha}{AC_\alpha} - 1 \right] > -\frac{1}{nD'}$$

This says that the minimum of restricted average cost must move under the boundary created by the per-firm demand.[3]

---

[3] An interesting addendum to this discussion is to address the question of what happens when trading of pollution permits is allowed.

The following graph (Figure 2) shows the model. It is useful to develop the idea of a competitive per firm demand curve. Divide the market quantity demanded by the number of firms in the industry. This gives a per firm demand curve. In long-run competitive equilibrium, this demand curve intersects the average cost curve of the firm at its minimum point.

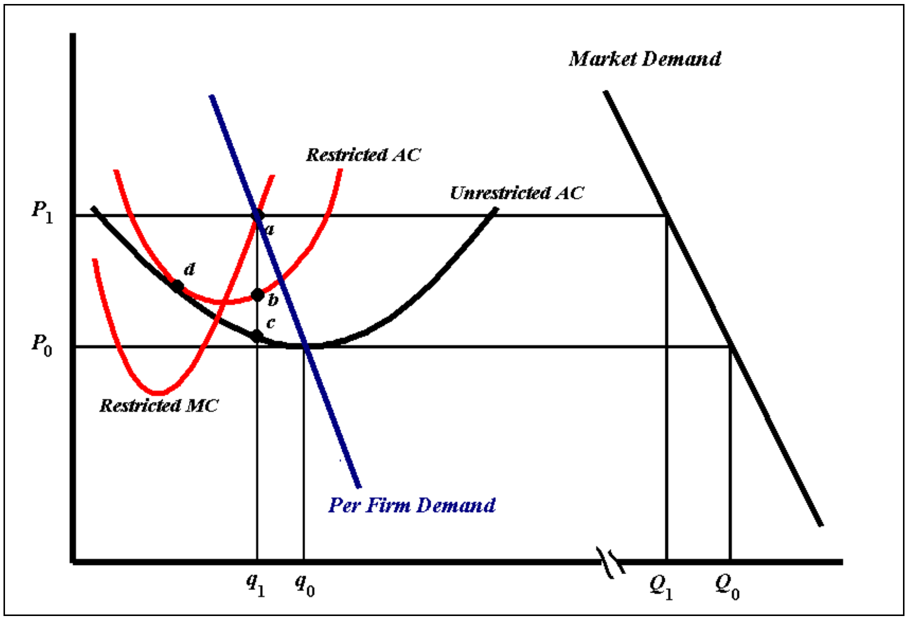


Figure 2

The input restriction moves the firm back along the unrestricted average cost curve. However, given the input restriction, the firm is able to compensate by using other inputs more intensively. The firm has a new marginal cost function that intersects the minimum of the restricted average cost. The new market equilibrium is determined where this new marginal cost function intersects the per-firm demand curve. The input restriction creates an output restriction of the amount [$Q_0$-$Q_1$] in the market which is [$q_0$-$q_1$] for the firm or $1/n^{th}$ of the market. The firms enjoy profits in the amount of (*a*-*b*) per unit. However, given this level of output restriction the inefficiency of (*b*-*c*) per unit is created.

## Acreage Restrictions versus Output Quotas

Notice that the analysis that we have just completed offers a straightforward answer to an old question: What is the difference between acreage restrictions in agriculture versus output quotas? Two things are clear from the graphical analysis that we have developed. First, an output quota, even if it is not transferable does not induce an inefficient mix of inputs given the firm's output, at least in the long run. Of course, if it is transferable, there is no production inefficiency. If it is not transferable, it does force firms to produce at an inefficiently small scale, but the input choice at that scale, in the long run, obeys the rule that the marginal produce ratio is equal to the input price ratio.

It is important to recognize the basis on which we can draw this conclusion. The quota forces the firm to reduce output, but it also allows the firm to produce that output without restriction on the use of any input. Hence, the firm will in the long run produce at point *c* in Figure

1. Point *c* is a point on the long run average cost curve. It is derived from the cost minimizing model given a target output of $q_1$. This cost is derived from the fact that the firm chooses inputs based on the FOC of that minimization model. These FOC say that input choice should follow the equi-marginal rule: the marginal product ratio between any two inputs should equal the price ratio between those inputs.

Second, acreage restrictions are inefficient. This is true because they allow farmers to substitute other resources for land in the production process. This is shown in Figure 2 as the distance (*b*-*c*). The firm expands output along the average cost function from point *d* to point *b*. Point *d* reflects the level of the input restriction at a point where there is no inefficient substitution into other inputs. However, the nature of the regulation allows for this inefficient substitution. As the firm expands output by using too little land, the inefficiency cost can be measured by the distance between the restricted and unrestricted average cost functions.

**Competitive Industry Equilibrium Effect Of An Input Price Change**

Now we move to the final analysis, that of the effect on the competitive market equilibrium that results from a change in an input price. This seems like it should be a simple exercise, but as we have seen in the earlier cases of quotas, taxes, and input restrictions, there are many nuances involved in the competitive response. Here, again, we maintain our assumption that the competitive market is comprised of identical firms each basing its production and planning decisions on a U-shaped long-run average cost function.

If AC is U-shaped we can focus on minimum AC as a point of reference for the competitive market equilibrium. At the quantity associated with min AC, AC($q$*)=MC($q$*). In other words, the competitive market equilibrium is defined by intersection of marginal and average cost. The star notation implies an equilibrium assumption, here that average and marginal cost are only equal at one point. Using the cost function implied by the optimal input demand functions, $x_i^*(w_1, \ldots, w_n, q)$, can write:

$$MC(w_1,\ldots,w_m,q^*) = AC(w_1,\ldots,w_m,q^*) \qquad 1$$

where $q^*$ is the output level that solves this equation.[4]

This formulation is used to determine how the long-run competitive equilibrium output level of the firm changes as an input price changes. We do this by differentiating with respect to $w_i$. This gives,

$$\frac{\partial q^*}{\partial w_i} = \frac{\dfrac{\partial AC}{\partial w_i} - \dfrac{\partial MC}{\partial w_i}}{\dfrac{\partial MC}{\partial q}} \qquad 2$$

Equation (3) allows for three different cases in the event of an input price change. In one case, MC goes up by more than AC. In this case $q$* shifts to the left. The other possibilities are that

[4] In this sense, $q^*$ is a function of the remaining parameters in the problem, the input prices, i.e., $q^*(w_1,\ldots,w_m)$.

MC cost goes up less than AC or that MC actually falls as an input price increases. In either of these cases, $q^*$ goes up.

To get a clearer picture of this, substitute for $\partial AC / \partial w_i$ and $\partial MC / \partial w_i$ using the envelope theorem.

$$\frac{\partial q^*}{\partial w_i} = \frac{\frac{x_i^*}{q} - \frac{\partial x_i^*}{\partial q}}{\frac{\partial MC}{\partial q}} = \frac{\frac{x_i^*}{q}\left[1 - \varepsilon_{iq}\right]}{\frac{\partial MC}{\partial q}} \qquad 3$$

This result says that the effect on the competitive equilibrium output level of the firm that is associated with an input price change depends on the output elasticity of the input. If the input is superior, then the output level of the firm will fall as the input price increases. Otherwise the equilibrium output level will rise.

If we go back to Alchian's principles on cost, we normally think of capital as a superior input. Average cost falls as we increase plant size because the productivity of big machines increases faster than their cost. It makes sense, then, that if the price of capital increases, the U-shaped average cost curve shifts up and to the left—economies of scale have become more expensive. On the other hand, for inputs that are not superior, inputs that do not contribute to economies of scale, inputs that have output elasticities less than one or inputs that are actually inferior, average cost shifts up and to the right—economies of scale have become *relatively* cheaper.

The model describes the movement in the quantity associated with minimum average cost. An increase in an input price will cause average cost to shift up and it may cause the minimum of average cost to shift to the right or to the left. Costs are higher and the optimal sized firm may be bigger or smaller.

Because average cost is higher, market price has to go up in the long run.[5] This means that market output must decrease. The question is, What happens to the number of firms in the market. There are two possibilities: (1) If the optimal firm size grows while market output is shrinking, then the number of firms must decline. (2) If the optimal firm size falls, the effect on the number of firms in the market depends on the relative sizes of these two effects. If the optimal firm size declines in percentage terms more than the reduction in market demand, then the number of firms in the market could actually increase as the industry shrinks. Conversely, if an input price declines, but this results in an increase in the optimal firm size, the number of firms could decrease even though the market is growing.[6]

It is fair at this juncture to return to the case of the input restriction. Recall that we were interested in the situation where the government restricted the use of an input but then allowed the restricted use to be transferable. Examples of such government policy are environmental pollution constraints and agricultural acreage restrictions. In the pollution case, the federal government mandated pollution restrictions on things like $SO_2$ emissions in 1970. At first these were hard and fast controls. However, over time they became transferable emission rights. Similarly, there have been numerous cases of acreage restriction that became transferable, though sometimes with limitations on geographical mobility.

[5] Even in the short run, if marginal cost is higher, market price must increase.

[6] Arguably this latter effect is at play in the computer industry as the cost of computing power has declined since the invention of the PC.

When input restrictions become transferable, the effect is exactly like an input whose supply has become perfectly inelastic. The initial right holders become input suppliers (possibly to themselves as producers).

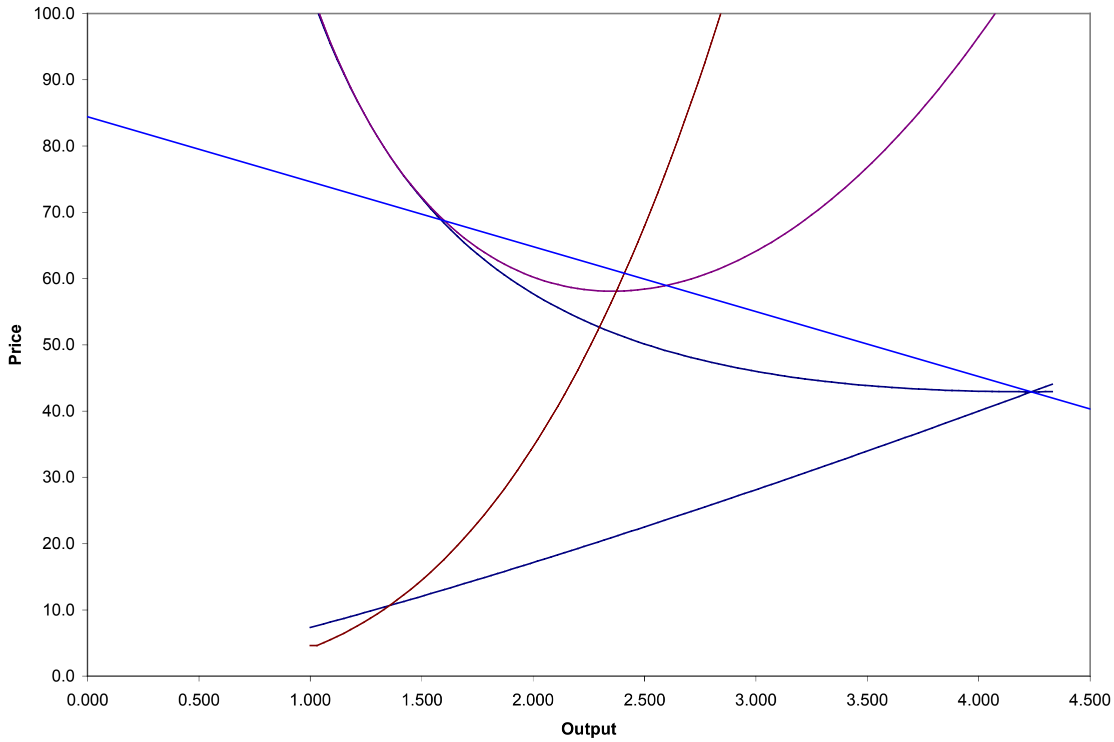


Figure 3

Figure 3 below shows the input restriction similar to Figure 2 above.[7] The per firm demand curve cuts average cost at its minimum and shows a zero-profit competitive equilibrium. Government regulation then reduces the use of an input which shifts this restricted average cost along the bowl of the unrestricted average cost function. Note that the restricted level of the input must have been optimal at some output level, which is generally true so long as the production function is continuous. In this example, the assumption is that restricted AC moves to the left along unrestricted average cost implying that the output elasticity of the restricted input is positive.

The competitive equilibrium given the restricted input usage is defined by the intersection of the per firm demand curve and the restricted marginal cost function. The assumption is that government regulation restricts entry into the market, which government must if the total industry usage of the input is to be restricted. That is, if regulation were to force individual firms to use less of a resource, but not restrict the total number of firms, the excess profits shown in Figure 3 would be dissipated by the entry of new firms, and total industry usage of the input would not change (substantially).

[7] These graphs are generated using a two-factor diminishing returns C-D production function with a fixed cost. One input is held fixed at an arbitrary level. Cost in this restricted cased is the cost of the fixed level of the input plus the cost of the variable input necessary (based on the production function) to achieve the target output, plus the fixed cost.

As discussed above, an example of an input restriction is pollution. When Congress passed the Clean Air Act in 1970, the regulations called for restrictions on the amount of pollution that firms could emit. Initially, the restrictions were specific controls at the point of source of the emissions. However, over time the regulations became restrictions on the actual amount of pollution that could be issued. As regulation evolved in this way, it was only natural that the restricted levels of the pollution could be traded. That is the input, albeit restricted at the industry level, became transferable. Other regulations have evolved similarly. Acreage restrictions in agriculture typically started out as restrictions on the use of land by each farm. Over time these became rights to plant a certain number of acres. Then these acreage rights became transferable.

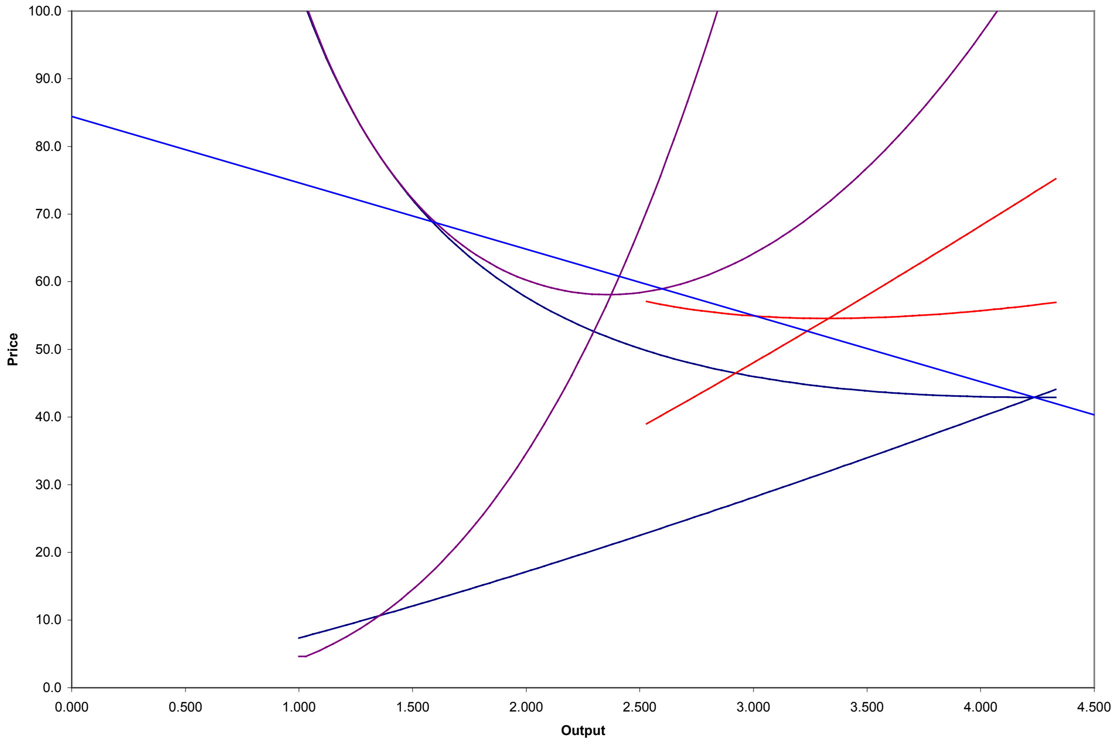


Figure 4

Figure 4 shows this initial situation when the market for the restricted input first opens. In order for a market to develop it must be true that the firm buying the restricted units can pay an amount at least as large as the firm initially endowed with them is earning by using them itself. This is the case based on the situation shown in Figure 4. A firm buying extra restricted units would have marginal and average costs that lie above the unrestricted cost curves and to the right of the restricted average and marginal cost curves. It could buy restricted units at a price that makes firms operating with restricted use of the input indifferent, produce at the market equilibrium price determined by the restricted MC, and make a profit.

Notice in Figure 4, the difference between the effect of the input restriction, which moves the restricted average cost up along the bowl of unrestricted average cost, and the effect of the input price increase for the firm that buys restricted units. The firm buying the restricted units sees the industry-wide input restriction as an increase in the input price. This shifts the marginal and average cost curves in the standard fashion, in this case, up and to the left. While it is not dramatically indicated in the graph, the price at which per firm demand intersects restricted MC is above the minimum of the average cost for the firm buying the restricted units.

Figure 4 shows us that a market in the restricted units will develop. However, it does not show an equilibrium. Two factors must adjust. First, there are excess profits for firms buying the restricted units, but more importantly, the number of firms in the industry must adjust. Because there is only a given amount of the restricted input, as firms acquire more of it than the initial restricted endowment, the number of firms in the industry must decline. Graphically, as the number of firms declines, the per firm demand curve pivots (on the choke price) outward.

Because the market for the restricted units are initially viable, the marginal cost of the firms acquiring additional units of the restricted input becomes the supply curve for the industry, and that average cost curve defines the market equilibrium. Market price can go up or down depending on the extent to which per firm output expands relative to the reduction in the number of firms. Based on the cost functions used in these illustrations, market price goes up. As you can see in Figure 5, the intersection of the original per firm demand curve and the restricted MC is below the intersection of the per firm demand when the restricted input is transferable and the MC of firms that buy restricted units. Also note that the price at this point is equal to average cost. The price of the restricted input rises to the point where profits are zero.

In some ways, it seems odd that market price when the restricted input is transferable can be higher than when it is not transferable. That is, consumers are made worse off by transferability. However, the harm caused to consumers is offset by the gains in production efficiency. When the restricted input is used in the initially allocated amounts, too much of the other resources are used along with it. So, for example, when the EPA initially assigned pollution restrictions, firms spent too much to clean up each and every little pollution source. It was and is more efficient to aggregate the pollution rights, reduce the number of firms and the amount of other resources that are spent producing goods.

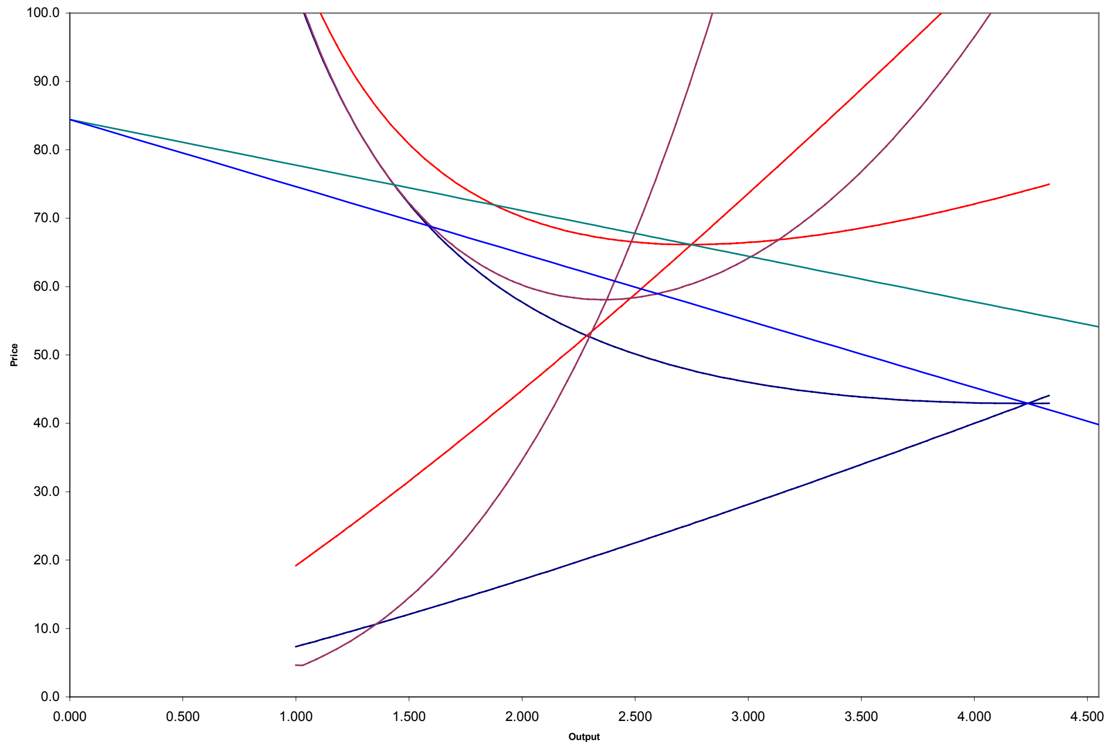
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Price
0.000
0.500
1.000
1.500
2.000
2.500
3.000
3.500
4.000
4.500
Output

Figure 5