

COMPETITIVE EQUILIBRIUM INPUT DEMAND¹

The simplest way to address this question is to model the industry as being characterized by constant returns to scale. When we assume that the industry is characterized by constant returns to scale, we are implicitly assuming that firm size is indeterminate. There could be only one firm in the industry or an infinite number of infinitely small firms: There are no economies of scale.

The assumption of constant returns to scale means that the industry will be competitive and that the cost minimizing mix of inputs will be chosen for any given level of industry output. For instance, if industry output is all produced by a single firm, because of potential competition owing to the constant returns nature of the production process, it is forced to set price equal to average cost. That is, economic profits are forced to zero.

Graphically this can be depicted as a market with a downward sloping demand and a flat supply curve. The intersection of demand and supply determines market output, which is the output of the single but competitive firm.

Because the firm is forced by potential competition to operate where price is equal to average cost, we can model the behavior of the firm and industry as that of a cost minimizer for any given level of output. In other words, we omit explicit discussion of the profit maximizing behavior of the firm. Implicitly we are assuming that the firm is a profit maximizer. However, because of the assumption of constant returns to scale for the firm and industry and the competitive nature of the industry implied by this assumption, price will always equal average cost as well as marginal cost. The assumption of constant returns to scale simplifies the analysis. (This assumption is employed in another analysis, that of the labor market for athletes, in a note appended to this lecture.)

The standard cost minimization model can be written as follows. We use capital letter notation for Q and X_i to indicate that the input and output levels are industry values. In the Lagrangian formulation, we have:

$$\min_{\{X_1, X_2\}} C = w_1 X_1 + w_2 X_2 + \mu [Q - f(X_1, X_2)] \quad 1$$

The FOC and SSOC are familiar to us, as are the behavioral functions that follow from the implicit function theorem:

$$X_i^* = X_i^*(w_1, w_2, Q), \text{ for } i = 1 \ \& \ 2 \quad 2$$

and

$$\mu^* = \mu^*(w_1, w_2) \quad 3$$

where the optimized value of the Lagrangian multiplier is in this case, as we know from the envelope theorem, marginal cost. Note that by assumption of constant returns to scale, marginal cost is not a function of output. Also from the envelope theorem we can identify average cost which is:

¹ Silberberg (1990) Section 8.9; Silberberg (2000) Section 8.9; Layard & Walters (1978) Ch 9; Nicholson (1998) Ch 21; Nicholson (2000) Ch 14; Varian (1992) Ch18.

$$AC(w_1, w_2) = \frac{C^*}{Q} = \frac{w_1 X_1^*(.) + w_2 X_2^*(.)}{Q} \quad 4$$

Here again, average cost is not a function of quantity because of the assumption of constant returns to scale.

The zero-profits market equilibrium condition determines the equilibrium value of Q , which we can write as:

$$Q^* = D(p) = D(AC) = D(\mu^*) \quad 5$$

where $D(p)$ is the (inverse) market demand curve, market price is equal to average cost, and marginal and average cost are equal because of constant returns to scale.

Because the market equilibrium condition is simultaneously determined along with the input demand curves, we must rewrite eqt (2) to include the expression for Q^* . I like to use double star notation to reflect the fact that the input demand functions are simultaneously satisfying the FOC for cost minimization and the output condition for market equilibrium. The industry equilibrium input demand is the input demand from our cost minimization model equated at the industry equilibrium level of output. That is,

$$X_i^{**} = X_i^*(w_1, w_2, Q^*), \text{ for } i = 1 \ \& \ 2 \quad 6$$

The Slope of Industry Demand

Equation (6) defines the industry demand curve for an input as the output constant demand curve shifts due to changes in industry output. Figure 1 highlights the idea.

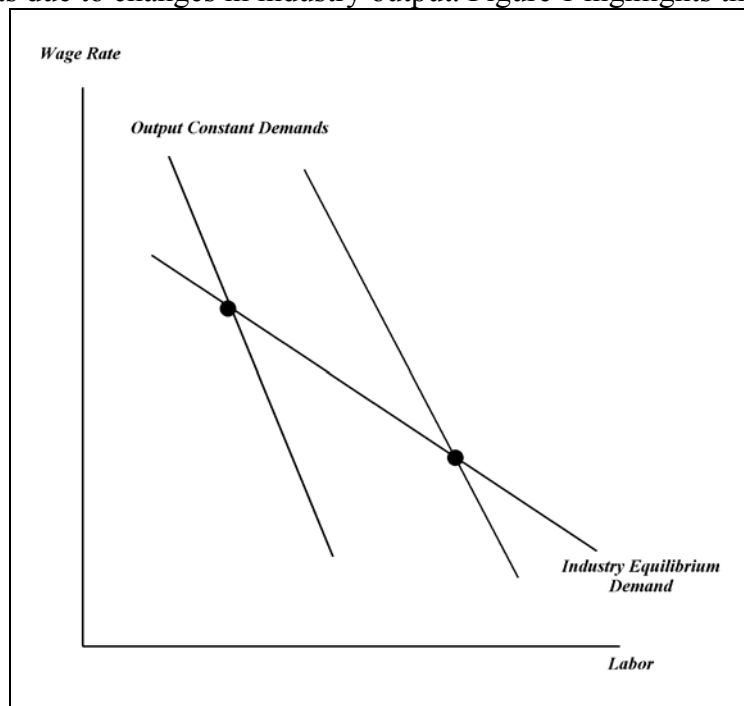


Figure 1

In this analysis, the output constant demand can be thought of as a short-run demand for an input and the demand given in eqt. (6) can be labeled long-run demand. The short/long run taxonomy is not straightforward in the CRS framework, but it is quite appropriate as we will see later on in these notes. As a consequence, you may substitute SR for the single star and LR for the double star subscripts.

The slope of the industry equilibrium input demand is composed of two parts. These can be shown by differentiating eqt (6), from which we get the following:

$$\left. \frac{\partial X_i^{**}}{\partial w_i} \right|_{Q^*} = \left. \frac{\partial X_i^*}{\partial w_i} \right|_{\bar{Q}} + \frac{\partial X_i^*}{\partial Q} \frac{\partial Q^*}{\partial w_i} \quad 7$$

The first term on the right-hand side of eqt (7) is always negative in the case the cost minimizing firm; it is the pure substitution effect. It is the movement along an isoquant in two-input space.

The second term on the right-hand side of eqt (7) is the output effect of the input demand function. It is always positive for homogeneous production functions. The last term is the market equilibrium output effect caused by an input price change. From eqt (5) we know:

$$\frac{\partial Q^*}{\partial w_i} = D'(p) \frac{\partial AC}{\partial w_i} \quad 8$$

This term is always negative because market demand is downward sloping.

From the envelope theorem and eqt (4) we know that

$$\frac{\partial AC}{\partial w_i} = \frac{\partial C^*}{\partial w_i} \frac{1}{Q} = \frac{X_i^*}{Q} \quad 9$$

Thus, we can write:

$$\frac{\partial X_i^{**}}{\partial w_i} = \frac{\partial X_i^*}{\partial w_i} + \frac{\partial X_i^*}{\partial Q} \frac{X_i^*}{Q} D' \quad 10$$

As it stands eqt. (10) is not very informative. It can be rewritten in elasticity form which makes it more appealing. First, multiply through by w_j/X_i ; second, multiply last expression by Q/Q and by p/p , where p is market price. These manipulations result in the following expression:

$$\varepsilon_{ii}^{**} = \varepsilon_{ii}^* + \varepsilon_{iQ}^* S_i \eta \quad 11$$

where the ε 's are input demand elasticity, S_j is the cost share of the i^{th} input, and η is price elasticity of market demand. Here, again, the star notation reflects the different levels of analysis. A single star is the demand curve assuming cost minimization subject to an output target. The double star is the competitive input demand assuming that the output market is in equilibrium.

Equation (11) can be further simplified by noting that ε_{iQ} , which is the output elasticity of the input, is equal to one in the case of constant returns to scale.² Hence, we have:

$$\varepsilon_{ii}^{**} = \varepsilon_{ii}^* + S_i \eta \quad 12$$

Equation (12) identifies a fundamental result for industry input demand functions. The elasticity of industry equilibrium input demand is equal to the output constant demand plus the cost share of the input times the price elasticity in the output market. The first term on the right-hand side of (12) is the pure substitution effect while the second term is the output effect. The pure substitution effect is always negative and the output effect is always negative because output demand is downward sloping. Hence, industry equilibrium input demand is always more elastic than output constant input demand as is shown in Figure 1.

We can go one step further from equation (12) by applying the concept of the *elasticity of substitution* [see next section].³ The elasticity of substitution is equal to the cross-price elasticity of demand divided by the cost share of the other input. That is,

$$\sigma = \frac{\varepsilon_{12}^*}{S_2} \quad 13$$

where σ is the elasticity of substitution. Since input demands are homogeneous of degree zero in prices, own-price elasticity is equal to the negative of cross-price elasticity. Hence, we can rewrite equation (12) as:

$$\varepsilon_{11}^{**} = -(1 - S_1)\sigma + S_1 \eta \quad 14$$

I have used explicit input numbers in the last two equations because some care must be used with the formula in equation (13). This formula is only true for the two input case, which is all right for constant returns production. We can think of input one as a unique input and input 2 as everything else. Even so, the elasticity of substitution in this case must be thought of as the substitutability of one for everything else. (See section on σ below.)

Equation (14) may be familiar from other presentations.⁴ It says that the elasticity of demand along the demand curve defined by the competitive adjustment is a function of three things: the elasticity of substitution, output demand elasticity, and the input's cost share. Notice that elasticity increases with cost share (Marshall's 2nd law) if output demand elasticity is bigger than the elasticity of substitution.

It is useful to consider equation (12) for the special case of the Cobb-Douglas production function. In the Cobb-Douglas case, own-price elasticity is equal to the negative of the quantity one minus the input's cost share (which itself is equal to the input's technical coefficient, i.e., the input elasticity in the production process). That is,

² At least for homogeneous production processes. It is possible to have non-homogeneous, constant returns to scale production processes, but these are necessarily idiosyncratic.

³ Silberberg (2000) circa p. 242.

⁴ Layard and Walters, p. 267.

$$\varepsilon_{ii}^* = -(1 - S_i)$$

and where $Q = K^\alpha L^\beta$, $\varepsilon_{KK}^* = -\beta$ and $\varepsilon_{LL}^* = -\alpha$. The industry equilibrium price elasticity is then

$$\varepsilon_{ii}^{**} = -(1 - S_i) + S_i \eta = S_i(1 + \eta) - 1$$

Obviously, if output demand is unit elastic, then input equilibrium demand is similarly so. If output demand is inelastic, then so is input demand. In this case η is less than one, and the first term is positive which makes elasticity less than one in absolute value. The bottom line is that for the Cobb-Douglas case, the input industry equilibrium demand elasticity follows the output demand elasticity.

A Note on Elasticity of Substitution

The elasticity of substitution (σ) has come into play in this lecture and some discussion of it is in order. The elasticity of substitution is the shape of the isoquant. The flatter is the isoquant, the greater is σ . The more "L" shaped, the smaller is σ . In this sense, the elasticity of substitution is aptly "defined" by the expression given by equation (13):

$$\sigma = \frac{\varepsilon_{12}^*}{S_2}$$

because the cross price elasticity is the economic representation of the shape of the isoquant. As the price of input two changes, holding input price one constant, the flatter is the isoquant, the more one is substituted for two. While this expression can be derived, I think it may be best to think of it as the definition of the elasticity of substitution.

The elasticity of substitution is symmetric because of the symmetry of the cross price effects.

$$\frac{\partial x_1^*}{\partial w_2} = \frac{\partial x_2^*}{\partial w_1}$$

In elasticities, we have:

$$\frac{\partial x_1^*}{\partial w_2} \frac{w_2}{x_1^*} \cdot \frac{x_1^* w_1}{C^*} = \frac{\partial x_2^*}{\partial w_1} \frac{w_1}{x_2^*} \cdot \frac{x_2^* w_2}{C^*}$$

which then gives:

$$\frac{\varepsilon_{12}^*}{S_2} = \frac{\varepsilon_{21}^*}{S_1} = \sigma$$

Also, in the two input case:

$$\varepsilon_{11}^* + \varepsilon_{12}^* = 0$$

so:

$$\varepsilon_{11}^* = -(1 - S_1) \cdot \sigma$$

The problem with the concept of the elasticity of substitution is that it can only be defined in terms of one input and all other inputs taken as a composite. In the n input case, if the

price of input one changes the elasticity of substitution defined between one and the composite of all other inputs assumes that all other inputs change by the same proportion.⁵ In practice this may or may not be true and defining σ based on this assumption is restrictive. However, thinking about the problem on the flip side is not so bad. The cross price elasticity of one with respect to the composite two tells us how much one changes if a price index of all other inputs changed. Thus, the elasticity of substitution is the substitutability of one for the general inflation of the prices of all other inputs. This has some intuition.

Inputs as Substitutes and Complements

Sometimes it is interesting to know if two inputs are substitutes or complements. For instance, it is often argued that unions favor minimum wage legislation not because of an altruistic spirit of camaraderie with unskilled labor but because skilled and unskilled labor are substitutes in the production process. If such is the case (not just as a technical characteristic of the production function, but as a revealed fact in the labor market) then by raising the minimum wage, the demand for skilled labor will increase.

The notion of substitutes and complements is defined in terms of the cross input price effects on the behavior of the firm. The cross pure effects in the output constant case are always positive in the two input model because it is a movement along the isoquant; the two inputs must be substitutes. However, when we allow the industry to adjust its output in response to the input price change, we get an output-weighted cross price effect that can go either way. Inputs can be substitutes or complements.

Following an outline similar to the development of the own price elasticity, we differentiate eqt (6):

$$\left. \frac{\partial X_i^{**}}{\partial w_j} \right|_{Q^*} = \left. \frac{\partial X_i^*}{\partial w_j} \right|_{\bar{Q}} + \frac{\partial X_i^*}{\partial Q} \frac{\partial Q^*}{\partial w_j} \quad 15$$

The first term on the right-hand side of eqt (15) is always positive in the case of the two input cost minimizing firm; it is the partial derivative of input demand w.r.t. the cross price and represents a movement along an isoquant because it is evaluated holding firm output constant. The second term on the right-hand side of eqt (15) is the output effect of the input demand function. It is always positive for homogeneous production functions. The last term is the market equilibrium output effect caused by an input price change. From eqt (5) we know:

$$\frac{\partial Q^*}{\partial w_j} = D'(p) \frac{\partial AC}{\partial w_j} \quad 16$$

This term is always negative because market demand is downward sloping. From the envelope theorem and eqt (4) we know that

$$\frac{\partial AC}{\partial w_j} = \frac{\partial C^*}{\partial w_j} \frac{1}{Q} = \frac{X_j^*}{Q} \quad 17$$

Thus, we can write:

⁵ It is very similar to the problem of separability in consumer goods.

$$\frac{\partial X_i^{**}}{\partial w_j} = \frac{\partial X_i^*}{\partial w_j} + \frac{\partial X_i^*}{\partial Q} \frac{X_j^*}{Q^*} D' \quad 18$$

The output effect of input demand is multiplied by the output effect of the change in market price. Because of the interactive nature of the two output effects, eqt (18) can be negative or positive.

Again, eqt. (18) can be rewritten in elasticity form to make it more revealing. This gives:

$$\varepsilon_{ij}^{**} = \varepsilon_{ij}^* + S_j \eta \quad 19$$

based on the fact that the output elasticity of the input must be one in the case constant returns to scale for homogeneous production processes. Further substitution from equation (13) gives:

$$\varepsilon_{21}^{**} = S_1 \cdot (\eta + \sigma) \quad 20$$

Equation (20) is the basic result. The sign of the competitive equilibrium cross price elasticity is determined by the relative size of the industry output demand elasticity and the elasticity of substitution between the two inputs. Output demand elasticity is negative. The elasticity of substitution is defined as positive.

In the special case of the Cobb-Douglas function, the elasticity of substitution is equal to 1, so we can derive the following result:

- If market demand is inelastic then the inputs are substitutes; if market demand is in the elastic region, then the inputs are complements.

The Cobb-Douglas form highlights the general form of the answer. Inputs are more likely to be substitutes in industries with inelastic demand than in industries with elastic demand. The prediction is, then, in the context of labor union support for minimum wage legislation, we should expect unions in the industries with elastic demand not to favor minimum wage laws.

Long-Run Industry Supply

Consider the question: If all firms in an industry operate according to linearly homogeneous production functions, can long-run industry supply be positively sloped? The answer is yes, and it is based on the rule that the elasticity of long-run industry supply is determined by the elasticity of industry input supply, not the technical aspects of the production function. It is worth sketching this out in detail.

Figure 2 shows the case where industry supply in the short-run is characterized by constant returns to scale. Average and marginal cost are flat. At the industry output demand depicted by D_0 , the industry input demand is $x_1^*(Q_0)$. The intersection of input demand and input supply yields an equilibrium price of w_1^0 , which defines marginal and average cost and results in the output equilibrium of $\{P_0, Q_0\}$. When industry output demand shifts to D_1 , input demand shifts out also. This raises the price of input one along its positively sloping supply. The higher input price increases the average and marginal cost of output supply. The new equilibrium results where the input demand $x_1^*(Q_1)$ results in an input equilibrium price of w_1^1 that gives a marginal and average

cost of P_1 and an output demand of Q_1 . Because the supply of input two is flat, the increase in its demand does not affect input price and there is no feedback effect on output supply.

The industry output equilibria of $\{P_0, Q_0\}$ and $\{P_1, Q_1\}$ imply simultaneous equilibria in both the input and output market. This is true because input demand is a function of the quantity of output supplied and, simultaneously, input demand determines input price that determines the price of output and, hence, the quantity demanded. The long-run industry supply is the locus of these equilibrium points for different levels of output demand.

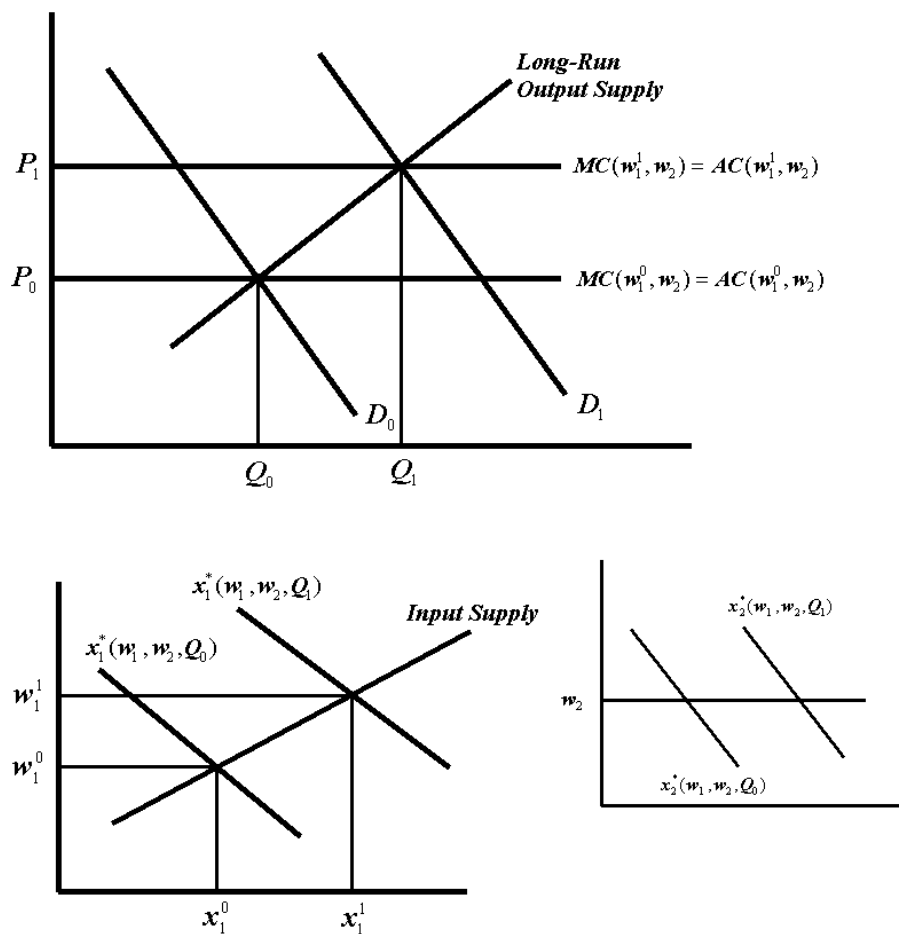


Figure 2

We can even go so far as to define the elasticity of output supply. In each input market, the equilibrium wage is defined by the condition that quantity demanded and supplied are equal. Quantity demanded (x_i^*) is a function of output as well as the input price. Quantity supplied of the input is only a function of the input price. The equilibrium input price is a function of output, $w_i^*(Q)$. The equilibrium wage elasticity with respect to output is given by:

$$\frac{\partial \ln w_i^*}{\partial \ln Q} = \frac{1}{\lambda_i - \varepsilon_{ii}^*}$$

where λ_i is the supply elasticity of input i . Note the demand elasticity that we used to define the equilibrium is the output constant form.

Industry output supply elasticity can then be defined as follows:

$$\frac{1}{\eta^s} = \frac{\partial \ln MC}{\partial \ln w_1} \frac{\partial \ln w_1^*}{\partial \ln Q} + \frac{\partial \ln MC}{\partial \ln w_2} \frac{\partial \ln w_2^*}{\partial \ln Q}$$

That is, a change in output will cause a change in input price, which then causes a change in MC .⁶ The effect is additive across the inputs.

The reaction of marginal cost to a change in input price is simply:

$$\frac{\partial \ln MC}{\partial \ln w_i} = \frac{\partial MC}{\partial w_i} \frac{w_i}{MC} = \frac{x_i^* w_i}{Q \cdot MC} = S_i$$

Note the use of the symmetry conditions. Also, this is still based on the assumption of CRS.

Thus, the industry supply elasticity is:

$$\frac{1}{\eta^s} = S_1 \cdot \frac{1}{\lambda_1 - \varepsilon_{11}^*} + S_2 \cdot \frac{1}{\lambda_2 - \varepsilon_{22}^*}$$

or

$$\frac{1}{\eta^s} = S_1 \cdot \frac{1}{\lambda_1 + (1 - S_1)\sigma} + (1 - S_1) \cdot \frac{1}{\lambda_2 + S_1\sigma}$$

where we substitute the elasticity of substitution expression for output constant input demand elasticity. Notice that if the supply of input 2 is perfectly elastic, that term vanishes and we can invert the other term simply.

Marshall's Last Law⁷

There is one more proposition that is usually professed in the discussion of industry input demand. It is:

Demand for a productive service will be more elastic, the more elastic is the supply of other productive services.⁸

Note that here, again, the law does not hold in the case where the elasticity of substitution is equal to the elasticity of output demand (absolute value). Marshall was imagining the case where the elasticity of substitution is less than output demand elasticity, and in this case as Stigler says the input owners with less than perfectly elastic supply bear, along with consumers, part of the burden of the increase in another input price. The increase in one input price causes market price

⁶ Notice that even though the chain of causation defining industry supply goes from output to price (unlike market demand), we still define supply elasticity in terms of the responsiveness of output to changes in price. At least I have done so in this derivation, following the lead of Nicholson's intermediate text.

⁷ Actually, it is his third law as given by Stigler, but it is the last one in the order that we treat them.

⁸ Stigler, Price Theory, p. 254.

to increase and market output to fall. In turn, this causes use of all inputs to fall. However, if one input is in less than perfectly elastic supply, its price falls which mitigates the output price increase.

Further discussion of this point and the long-run equilibrium is given at the end of this lecture.

CES Production Functions

Another specific production function that is widely referred to is the Constant Elasticity of Substitution form. It is given as:

$$q = (ax_1^\rho + (1-a)x_2^\rho)^{1/\rho}$$

Without going into great detail in the discussion of the CES form, let's just hit a few highlights. First, the elasticity of substitution is given by

$$\sigma = \frac{1}{1-\rho}$$

Also recognize that the CES production function is constant returns to scale. When each input is multiplied by a scalar, t , the scalar can be factored out with an exponent of 1. This is shown below:

$$q(t) = (a\{tx_1\}^\rho + (1-a)\{tx_2\}^\rho)^{1/\rho} = tq$$

Second, the elasticity of output with respect to an input can be derived as follows. First take the log of the production function and then the derivative of the log of output with respect to the level of the input:

$$\ln q = \frac{1}{\rho} \ln(ax_1^\rho + (1-a)x_2^\rho)$$

$$\frac{\partial \ln q}{\partial x_1} = \frac{1}{\rho} \frac{1}{ax_1^\rho + (1-a)x_2^\rho} \rho ax_1^{\rho-1}$$

$$\frac{\partial \ln q}{\partial x_1 / x_1} = \frac{ax_1^\rho}{ax_1^\rho + (1-a)x_2^\rho} = a \left(\frac{x_1}{q} \right)^\rho \quad 21$$

This last expression is the elasticity of output with respect to an input.

If we use the CES production function to describe the industry equilibrium as we did above with the Cobb-Douglas form, we go through the same steps. First, develop the model of cost minimization for a target level of output. Using the CES form, this is done most simply in the following way:

$$\min C = w_1x_1 + w_2x_2 + \mu[q^\rho - (a_1x_1^\rho - a_2x_2^\rho)]$$

where the target output level is monotonically transformed in the lagrangian by the exponent ρ . Note that $a_2 = (1 - a_1)$ and is substituted for notation convenience.

The first two first order conditions look like:

$$w_i = \mu a_i x_i^{\rho-1}$$

Solving for μ in these two equations gives:

$$\frac{w_1 x_1}{w_2 x_2} = \frac{a_1 x_1^\rho}{a_2 x_2^\rho}$$

Adding one to both sides of the equation allows us to add the denominator to the numerator on both sides. This gives:

$$\frac{C}{w_2 x_2} = \frac{a_1 x_1^\rho + a_2 x_2^\rho}{a_2 x_2^\rho}$$

Inverting we have:

$$\frac{w_2 x_2}{C} = \frac{a_2 x_2^\rho}{a_1 x_1^\rho + a_2 x_2^\rho} \quad 22$$

Notice that this says that the input's share is equal to the output elasticity with respect to an input derived in eqt. (21) above.

The CES cost and input functions can be identified by substituting from the production function in the denominator of (22).

$$\frac{w_2 x_2}{C} = \frac{a_2 x_2^\rho}{q^\rho}$$

Solve for x_2 ; repeat the process for x_1 ; and then substitute into the cost equation to get the cost function. The cost function looks like this:

$$C^* = q [a_1^{1/\rho} w_1^{-\rho/\rho} + a_2^{1/\rho} w_2^{-\rho/\rho}]^{1-\rho}$$

Remember that the derivative of the cost function with respect to an input price is x_i^* , which given an explicit functional form defines the input demand function.

$$x_1^* = q [a_1^{1/\rho} w_1^{-\rho/\rho} + a_2^{1/\rho} w_2^{-\rho/\rho}]^{-1/\rho} a_1^{1/\rho} w_1^{-1/\rho}$$

Taking the derivative with respect to own price and cross multiplying to get own price elasticity gives:

$$\frac{\partial x_1^*}{\partial w_1} \frac{w_1}{x_1^*} = -\frac{1}{1-\rho} \left[1 - \frac{\left(\frac{a_1}{w_1^\rho} \right)^{1/\rho}}{\left(\frac{a_1}{w_1^\rho} \right)^{1/\rho} + \left(\frac{a_2}{w_2^\rho} \right)^{1/\rho}} \right]$$

which is:⁹

$$\varepsilon_{11} = -\sigma(1 - S_1)$$

Summary & Applications

Here are some of the basic applications of production theory.

In the competitive equilibrium, price equals marginal cost. Hence, the percentage change in equilibrium output price that occurs when an input price changes is given by:

$$\% \Delta P = \left(\frac{\partial \ln MC}{\partial \ln w_i} \right) \% \Delta w_i$$

In the competitive equilibrium, marginal cost equals average cost, so the percentage change in marginal cost is equal to the percentage change in average cost. From the envelope theorem, we know that:

$$\frac{\partial AC}{\partial w_i} = \frac{X_i}{Q}$$

which means that:

$$\frac{\partial AC}{\partial w_i} \frac{w_i}{AC} = \frac{w_i X_i}{AC \cdot Q} = S_i$$

That is, for the case of Constant Returns to Scale, the elasticity of marginal and average cost with respect to an input price is always equal to the input's share of production cost. Thus, the predicted percentage change in market price due to a change in input price is cost share times the percentage change in the input price, i.e.,

$$\% \Delta P = S_i \cdot \% \Delta w_i$$

This is true for all CRS production functions regardless of the elasticity of substitution.¹⁰

⁹ I do not know how to derive this result from the expression for demand elasticity given above. The term in brackets looks like it is probably one minus cost share, but... The expression can be derived from the share equation directly.

¹⁰ For the Cobb-Douglas production function, marginal cost is

$$MC = \mu^* = \alpha^{-\alpha} (1 - \alpha)^{-(1-\alpha)} w_1^\alpha w_2^{(1-\alpha)}$$

which means that:

$$\frac{\partial \ln MC}{\partial \ln w_1} = \alpha$$

Furthermore, we know that α is input one's share of production cost.

In terms of input demand, the percentage change in input demand is the price elasticity of demand for the input times the percentage change in input price. There are two demand elasticities: short run, which holds market output constant, and long run where output changes to its new equilibrium value. Thus we can write:

$$\text{Short run: } \% \Delta X_i = \varepsilon_{ii}^* \cdot \% \Delta w_i$$

$$\text{Long run: } \% \Delta X_i = \varepsilon_{ii}^{**} \cdot \% \Delta w_i$$

Since the price elasticity of input demand depends on the shape of the isoquant, it is necessary to express price elasticity in terms of the elasticity of substitution. Thus, we have:

$$\varepsilon_{ii}^* = -(1 - S_i)\sigma$$

and

$$\varepsilon_{ii}^{**} = -(1 - S_i)\sigma + S_i\eta$$

where σ is the elasticity of substitution and η is the price elasticity of output demand.

Lastly, we have the input substitution effect and again there is a short run and long run phenomenon.

$$\text{Short run: } \% \Delta X_j = \varepsilon_{ji}^* \cdot \% \Delta w_i$$

$$\text{Long run: } \% \Delta X_j = \varepsilon_{ji}^{**} \cdot \% \Delta w_i$$

Here, we will treat input j as a composite of all other inputs but i . So, we have the two input case and the cross elasticity of j with respect to i from the homogeneity condition is given by:

$$\varepsilon_{ji}^* = -\varepsilon_{ij}^* = S_i\sigma$$

The long-run cross price elasticity is, then:

$$\varepsilon_{ji}^{**} = S_i(\sigma + \eta)$$

Let's apply these to the flower shop problem. Recall the problem:

Morgan's Flower Shop specializes in cut flower arrangements. It takes Sally about 15 minutes per arrangement. Sally is paid \$20 per hour. The flowers themselves cost around \$15 per arrangement. The arrangements sell for \$45. The flower business is highly competitive, and previous studies of this industry reveal that

1. the demand for cut flower arrangements is unit elastic, and
2. the industry is characterized by constant returns to scale; indeed, the Cobb-Douglas production function is a useful approximation of production at the firm level.

Assume that the wholesale cost of flowers that Sally has been using in each arrangement rises to \$20, from its previous level of \$15.

- A) What is your best estimate of the resulting increase in the retail price of flower arrangements?
- B) How and by how much will the amount of flowers in each arrangement change?
- C) How and by how much will the amount of wholesale flowers used in the industry change?
- D) How and by how much will the time that Sally spends on each arrangement change?
- E) How and by how much will the total number of flower arrangers like Sally change?

Applying our formula, we have:

A) $\% \Delta P = S_i \cdot \% \Delta w_i = \frac{1}{3} \cdot \frac{1}{3}$; or $\frac{1}{9}$ th of \$45 which is \$5. Note that if you solve this using the Cobb-Douglas function explicitly, the price change is \$4.5; error in part is the forecast error of the percentage change approximation of the change.

B) This is the short run own price elasticity effect: $\% \Delta X_i = -(1 - S_i) \cdot \% \Delta w_i = -\frac{2}{3} \cdot \frac{1}{3}$.

C) This is simply the long run own price elasticity:

$$\% \Delta X_i = \varepsilon_{ii}^{**} \cdot \% \Delta w_i = [-\sigma + S_i(\sigma + \eta)] \cdot \% \Delta w_i$$

For the C-D production, $\sigma = 1$, and with unitary output demand, the long run own price elasticity is $[-1]$. Thus, the reduction in the amount of flowers is

$$\% \Delta X_i = -\% \Delta w_i = -\frac{1}{3}$$

Sally is using two-ninths less flowers per arrangement. The other ninth reduction comes from the decline in the number of arrangements sold.

D) This is the short run cross price elasticity effect: $\% \Delta X_j = S_i \cdot \% \Delta w_i = \frac{1}{3} \cdot \frac{1}{3}$. That is Sally increases her time per flower arrangement from 15 minutes per arrangement to nearly 17 minutes.

E) This is the long run cross price elasticity effect. We know that in the long run, the change in the amount of time spent arranging flowers is zero:

$$\text{Long run: } \% \Delta X_j = (\sigma + \eta) \cdot \% \Delta w_i$$

because output demand elasticity in this problem is $[-1]$ and the elasticity of substitution is $[+1]$. Now, since Sally is spending $\frac{1}{9}$ th more time per arrangement, there must be a $\frac{1}{9}$ th reduction in the number of Sally's in the market.

CES Production with Positively Sloped Supply of a Resource:

This is a takeoff on the flower problem. Here I have assumed CES production with $\sigma = 0.333$. The CES production function is constant returns to scale (CRS). I initialize everything using the parameters given in the Morgan's Flower Shop problem. That is, the unit cost of labor is \$5, cut flowers are \$15, overhead is \$25, which gives a price of arrangements of \$45. All quantities are unitized. The shares give the initial linear parameters in the CES p.f.¹¹

I allow the supply of labor to vary in elasticity from $(1/2.5)$ to 10, which is the equivalent of the slope of the supply of labor varying from 2.5 to .1. Note that we usually and confusingly think about supply elasticity as bigger equals flatter.

The comparative static adjustment is to double the price of flowers. This is different from the original problem, but it simplifies the calculations. I simulate the model in SAS allowing the market equilibrium to be reached iteratively. One hundred iterations are plenty. The SAS code is shown at the bottom.¹

Doubling the price of flowers has two effects: the substitution of other inputs for flowers and an increase in the price of flower arrangements. The former causes the amount of other inputs to increase per unit of market output, but the latter causes the amount of market output to decline. This, in turn, causes the amount of all inputs to decline.

The following are a series of graphs that depict the results. The data were exported to Excel and the graphs were drawn using 'XY (Scatter)' graph type.

The first is the effect of the increase in flower cost on marginal cost. At the flattest supply of labor, MC (which is market price of output) increases from \$45 to nearly \$59. If we estimated the increase based on elasticities, we would forecast an increase in price to \$60. Partly the difference is the error of the forecast; partly the difference is the equilibrium change in the price of labor.

Notice that as the slope of the supply of labor becomes steeper, market price of flower arrangements increases less. This is the effect of Marshall's third law: as Stigler says, the steeper the supply of labor, the more it shares in the effect of the increase in the price of flowers, mitigating the effect of the flower price increase on consumers of flower arrangements.

This point is shown in more detail in the next two graphs. In Figure 2, we see that the industry use of labor falls the most when labor is supplied very elastically, and that if the supply of labor is more inelastic (i.e., steeper), industry use increases. This is because, as we see in Figure 3, the wage rate falls most when labor is supplied inelastically.

The increase in the price of flowers increases cost. This increases market price and reduces market output. The decline in market output is the dominant force on the demand for inputs, true because we have assumed that the elasticity of substitution is less than the absolute value of the elasticity of output demand. As the demands for all inputs fall, the decline for labor

¹¹ True because all inputs and output are initially set to unity.

causes its price to fall. This mitigates the increase in the price of output and the decline in market output.

We see in Figure 2, that when labor is supplied inelastically, there will be less flowers used per unit of output (i.e., flower arrangements). This is further emphasized in Figure 4 where we see that the labor to flowers ratio (in a unit of output) is increasing in the slope of the labor supply. This is also true for the labor to output ratio.

So at bottom what we know is this: In the case where the elasticity of substitution is less than the absolute elasticity of demand for output (the most reasonable case), an exogenous increase in the price of one input will cause:

- a) the price of output to increase and the demand for output to fall;
- b) the use of other inputs to increase per unit of output;
- c) the use of all inputs to fall in total;
- d) the most inelastically supplied inputs to increase the most per unit output and decline the least in total.

This last point is Marshall's third law: the industry-wide elasticity of demand for an input will be greatest when other inputs are most elastically supplied. Or, put differently, the industry-wide elasticity of demand for an input will be the lower the more inelastic is the supply of other inputs. This is shown in Figure 5.

Figure 1.

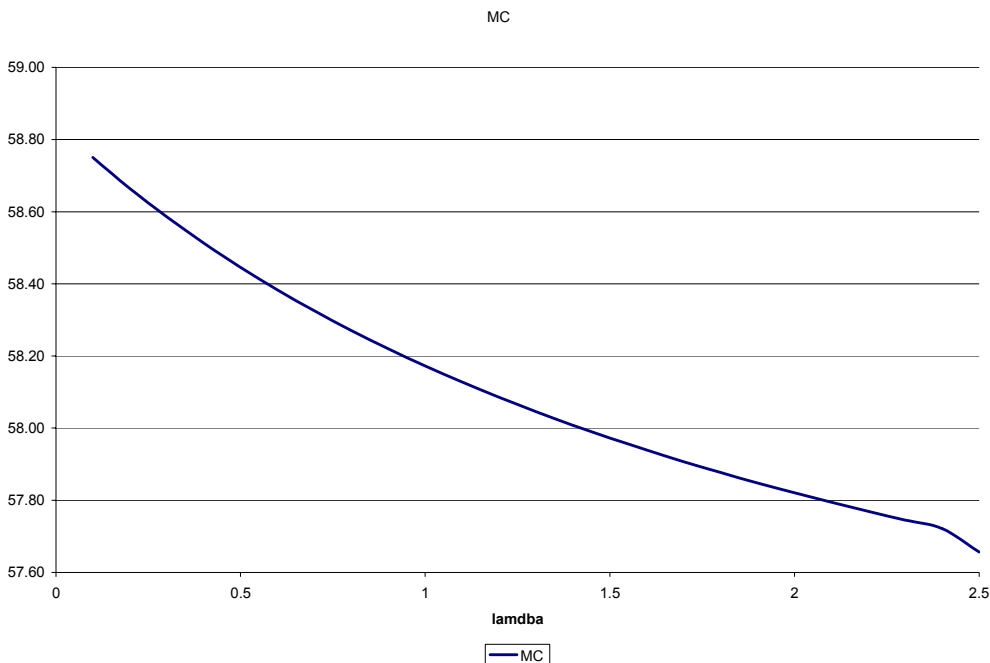


Figure 2.

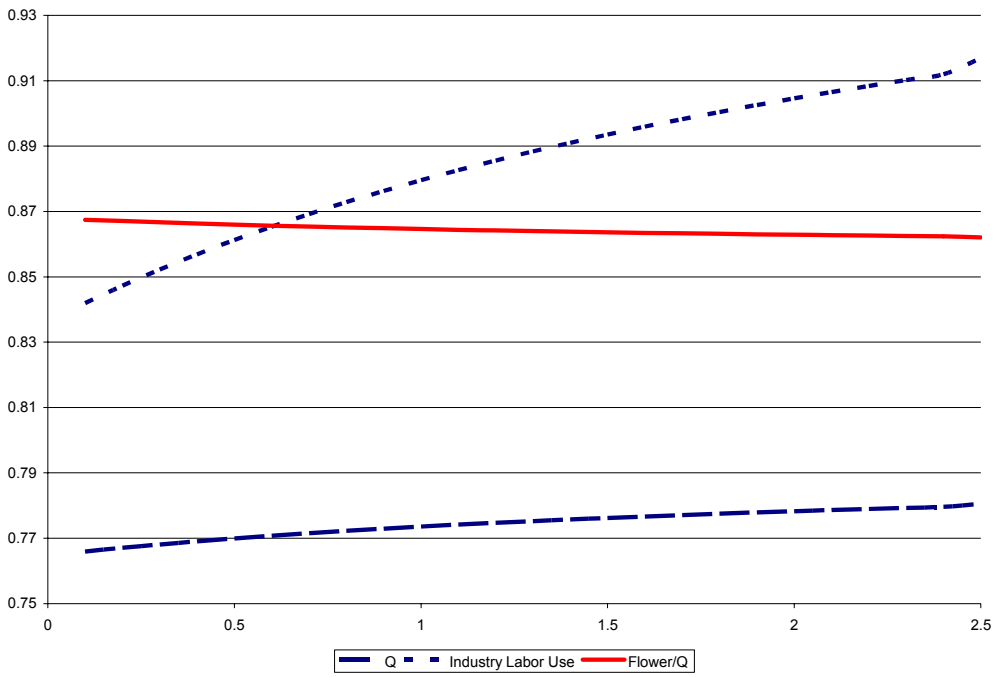


Figure 3.

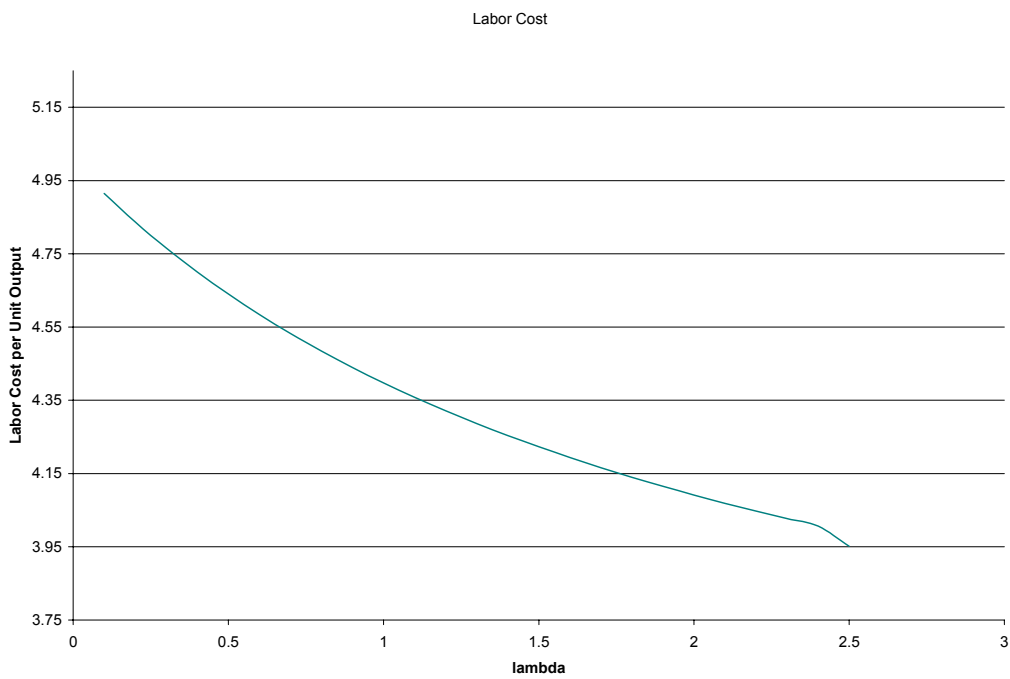


Figure 4.

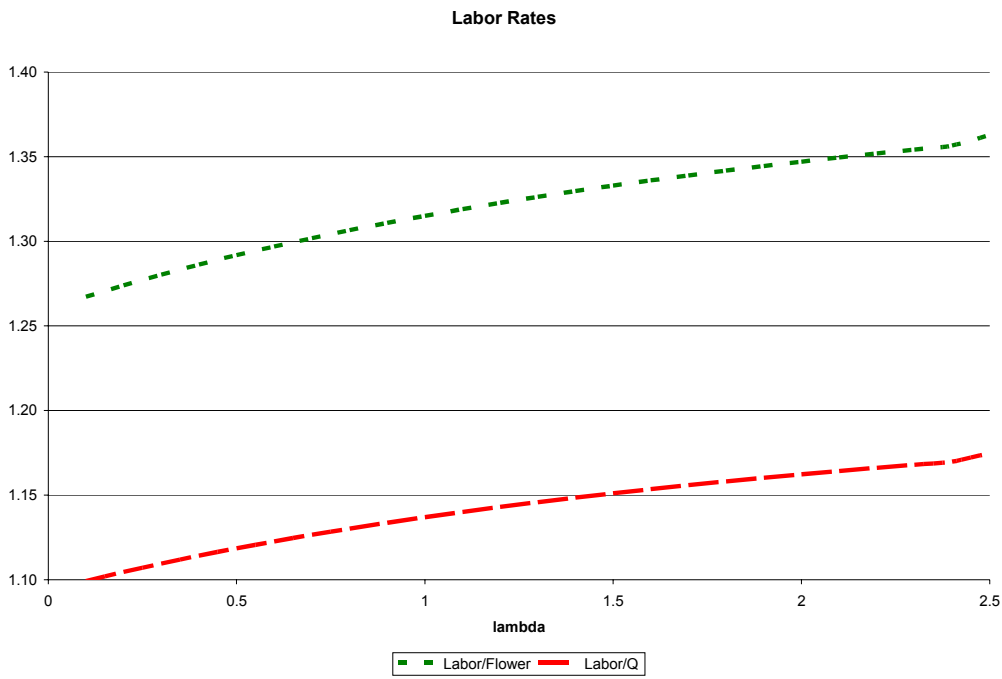
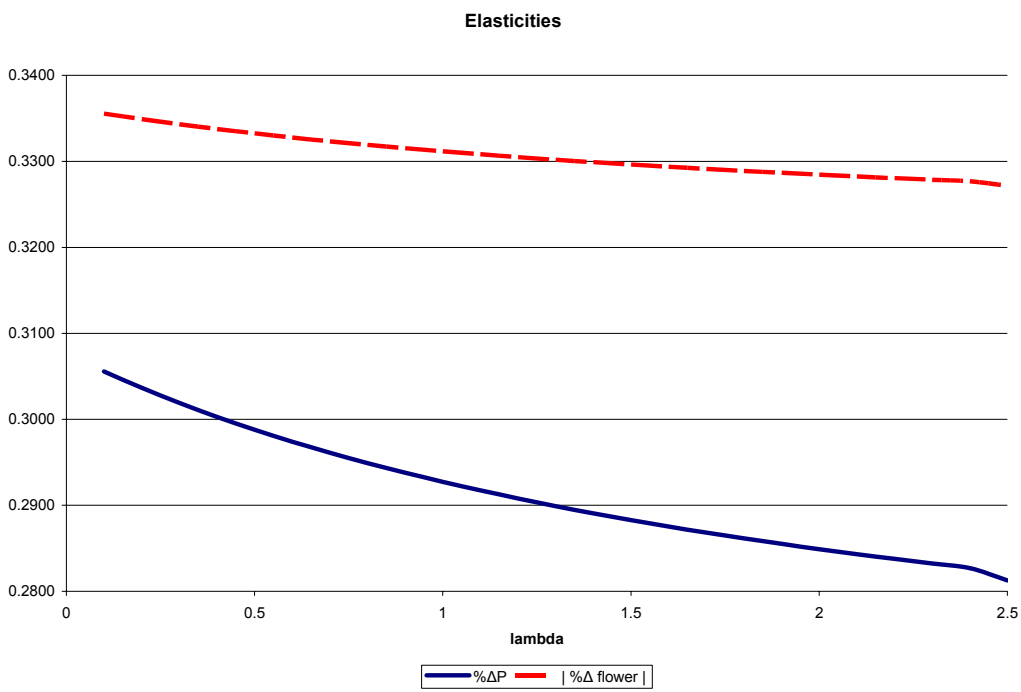


Figure 5.



Long-Run Supply

Using the same approach we can depict the long-run supply by allowing for exogenous changes in output demand. This program is also shown at the end.² The program maps out the demand by means of the iterative nature of convergence to an equilibrium after the shift in demand. The initial equilibrium is \$45. The increase in the demand for output causes the demand for inputs to increase. Labor supply is positively sloped. This causes the wage rate to increase, which causes MC to increase. The increase in MC causes the quantity demanded for output to fall. The process cycles until equilibrium convergence occurs.

Figure 6 shows the output market long-run supply and Figure 7 shows the input market equilibria. The demand curves for output in Figure 6 are indicated by the unconnected pattern of dots that fan out in a downward sloping line intersecting different points along the long-run supply. The supply curve itself is a “trend” line connecting the equilibrium values of each fanning pattern. The same is true in Figure 7.

Figure 6.

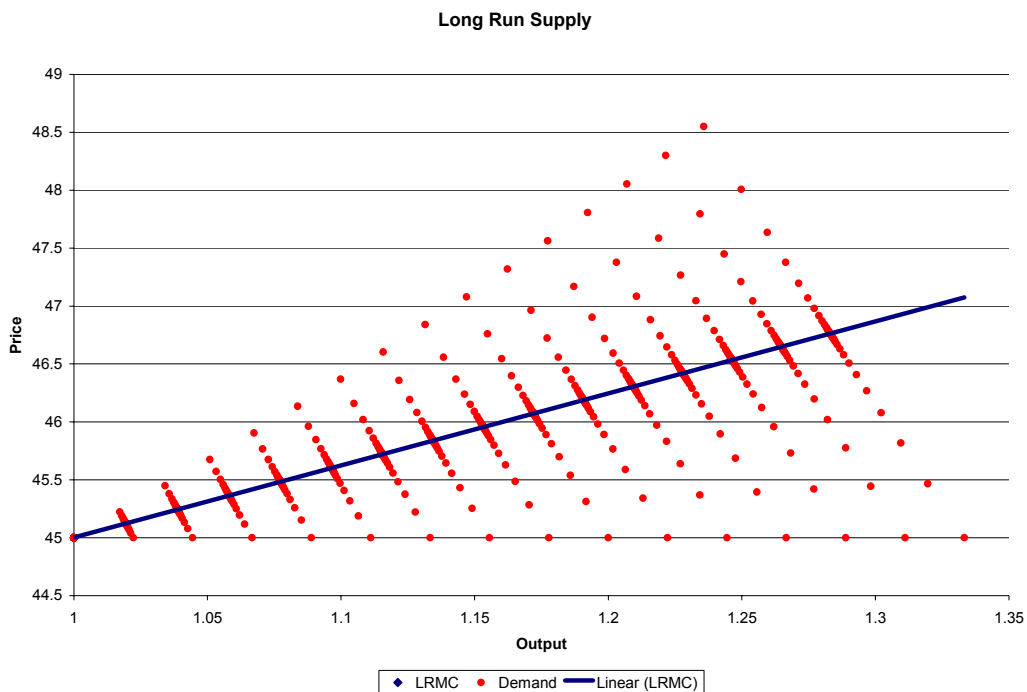
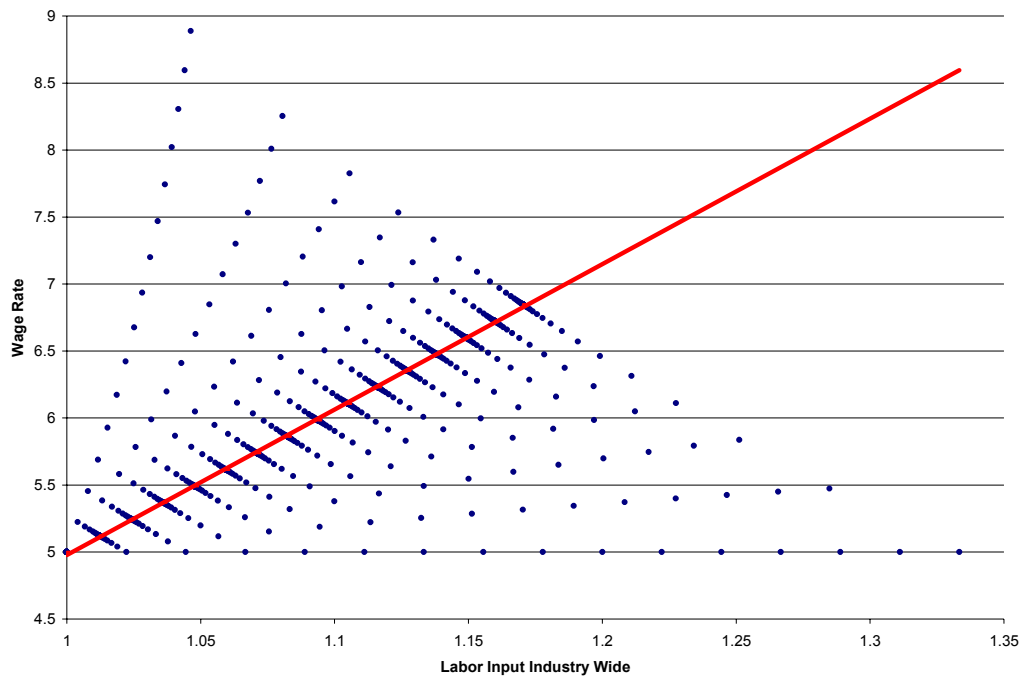


Figure 7.

**Putting it all Together:**

Let demand and supply be characterized by the constant elasticity forms:¹²

$$Q^D = DP^{-\beta}$$

and

$$Q^S = \Gamma P^\gamma$$

where the latter is the long-run supply of output. Thus, equilibrium price (designated by the star notation) is given by

$$P^* = \left(\frac{D}{\Gamma} \right)^{\frac{1}{\gamma + \beta}}$$

or in log form:

$$\ln P^* = \frac{1}{\gamma + \beta} (\ln D - \ln \Gamma)$$

So, we can write:

$$\frac{\partial \ln P^*}{\partial \ln \Gamma} = -\frac{1}{\gamma + \beta}$$

¹² The problem is set up using constant elasticity forms because they map most simply into the elasticity expressions for changes in input demand and output price. Linear demand and supply can be used as well with the appropriate conversion to elasticity form in the derivation.

which is the percentage rate of change in equilibrium price with respect to a shift in the long-run supply curve.

We know that the change in the supply price in percentage terms is the input's share of cost times the percentage change in the input price. Let this value be u . A shift in supply price of flowers is represented by a change in Γ in our supply function above. Rewriting the supply function:

$$u = \ln P = \frac{1}{\gamma} (\ln Q^S - \ln \Gamma)$$

The value of u from our last example is: $u = (100\%) \times (1/3) = 0.333$. That is, we let the price of cut flowers double, and the cost share of cut flowers is one third. The corresponding change in Γ is then $[-u \gamma]$ or $[-0.333 \gamma]$.

The problem now is to find the value of γ , which is the elasticity of output supply. This value is given by the formula that we derived above:

$$\gamma = \frac{\lambda_L + (1 - S_L)\sigma}{S_L}$$

where λ_L is the supply elasticity of labor, S_L is labor's share of cost, and σ is the elasticity of substitution. (Recall that we can invert this formula this way only when there is only one input with a non-flat supply.) In our problem, these terms take the values:

$$\gamma = \frac{.5 + (1 - .1111) \cdot .333}{.1111} \doteq 7$$

Now, we can estimate the effect of doubling the price of flowers on the equilibrium price of flower arrangements given the positively sloped supply of flower arrangements caused by the inelastic supply of labor. In general, we have:

$$d \ln P^* = \left(\frac{\partial \ln P^*}{\partial \ln \Gamma} \right) d \ln \Gamma = -\frac{1}{\gamma + \beta} d \ln \Gamma$$

Substituting for $d \ln \Gamma$:

$$d \ln P^* = u \cdot \frac{\gamma}{\gamma + \beta} = 0.333 \cdot \frac{\gamma}{\gamma + \beta} = 0.333 \cdot \frac{7}{8} = 0.29$$

This value is slightly off of the calculated value from the SAS program, which gives a value of 0.285. There are a number of explanations. First, there is always the problem of differential forecasts. This is error that they always teach in the first calc class. But in this problem we also have the problem that input shares are changing all along and the estimate given above uses the original input shares.

HAS “FREE-AGENCY” IN PROFESSIONAL SPORTS CAUSED TICKET PRICES TO GO UP?¹³

M.T. Maloney
Clemson University

Introduction

There is no cut and dried answer to the question of free agency in professional sports. What I do in this note is develop different models of the labor and output markets, and investigate the predictions that these models make. As a preview, we can show that the model of perfect price discrimination is identical to the model of perfect competition in its predictions about the price of output. Thus, if the reserve clause allowed professional teams to perfectly price discriminate and if free agency creates a perfectly competitive market, output price should be unchanged. On the other hand, free agency may not mean perfect competition, but rather results in a less comprehensive form of market power exercised by the owners over the players. In this event, there are predictable effects on output price. The note concludes with a comparison of the predictions of the different models in relation to the evidence as we know it.

Perfect Price Discrimination

Let's simplify the problem to the maximum extent. Assume that each team in the sports league operates according to a Cobb-Douglas, constant returns to scale production process [$q = K^\alpha L^\beta$]. Assume that all market areas are identical. Further, let's assume that the supply curve of players available to the whole league is given by the simple expression:

$$w_L = \lambda L ,$$

which comes from the player's decision about work and leisure. Assume all players are the same in productivity.

Under the reserve clause regime, each team is allocated a stock of players. There is no incentive to trade players because all players are identically productive. However, by means of trading players, owners can discover the supply prices of players. For instance, if Cleveland cuts the salary of player, Alomar, to the point that he will not play, Cleveland puts him on the trading block. Under the reserve clause, Cleveland has the right to stop Alomar from playing unless the team that hires him pays Cleveland what Cleveland demands in compensation. Hence, Cleveland will charge an amount that is \$1 less than the difference between Alomar's value to Cleveland and the price Cleveland offered him. If nobody takes him, Cleveland cuts its compensation another \$1, etc., until Alomar eventually signs with another club. The reserve clause offers owners a method of perfectly price discriminating against players.

Perfect price discrimination is identical to perfect competition in terms of the employment of inputs and in terms of the derived marginal cost of output. It is useful to demonstrate this result.

¹³ Frank, Jeff and Smith, Eric. "Seniority Seating at the Royal Opera House" in *Oxford Economic Papers* vol. 48, p 492-498. Cheung, Steven N. S. "Why are better seats underpriced"? in *Economic Inquiry* vol. 15, 1979. p 513-522

Under perfect price discrimination, the objective function of the owners, who can be treated as one economic agent, can be described as minimizing cost subject to an output objective. The cost of the player input is:

$$\int w_L dL = \int \lambda L dL = \frac{\lambda}{2} L^2$$

The objective function becomes:

$$\min C = w_K K + \frac{\lambda}{2} L^2 + \mu(q - K^\alpha L^\beta)$$

The demand curve of labor holding output constant is:

$$\hat{L} = \left(\frac{\hat{\mu} \cdot \beta \cdot q}{\lambda} \right)^{1/2}$$

where the important term is the marginal cost curve. It can be shown to be:

$$\hat{\mu} = q^{\frac{\beta}{2-\beta}} \lambda^{\frac{\beta}{2-\beta}} \beta^{-\frac{\beta}{2-\beta}} (\alpha^{-\alpha} w_K^\alpha)^{\frac{2}{2-\beta}}$$

Note that this marginal cost curve for output is not built on any assumption about relative ability of owners to monopolize the output market. While we are willing to assume that there is monopoly power in the output market, we will address that in good time. For now we have a marginal cost curve that is built simply on the ability of the league to perfectly monopsonize the supply of labor.

Perfect Price Discrimination Compared To Perfect Competition

It is useful to demonstrate that perfect monopsonization yields the same behavior and the same marginal cost (of output) function as competition. The competitive model is taken from the standard analysis:

$$\min C = w_K K + w_L L + \mu(q - K^\alpha L^\beta)$$

and the cost minimizing use of labor is:

$$L^* = \frac{\mu^* \cdot \beta \cdot q}{w_L}$$

where marginal cost is:

$$\mu^* = \alpha^{-\alpha} \beta^{-\beta} w_K^\alpha w_L^\beta$$

The input market equilibrium is competitively determined where the quantity of labor demanded is equal to the quantity supplied. Thus,

$$\frac{w_L}{\lambda} = L^s = L^* = \frac{\alpha^{-\alpha} \beta^{-\beta} w_K^\alpha w_L^\beta \beta q}{w_L}$$

Solving for the competitive equilibrium wage, which is a function of market output, gives:

$$w_L = [\lambda \alpha^{-\alpha} \beta^{1-\beta} w_K^\alpha q]^{\frac{1}{2-\beta}}$$

Substituting this competitive equilibrium wage back into the marginal cost curve and using double star notation to indicate that competitive marginal cost is based on (a) cost minimizing behavior on the part of the firms and (b) input market equilibrium, we have:

$$\mu^{**} = \alpha^{-\alpha} \beta^{-\beta} w_K^\alpha (\lambda \alpha^{-\alpha} \beta^{1-\beta} w_K^\alpha q)^{\frac{\beta}{2-\beta}}$$

It can be shown that:¹⁴

$$\hat{\mu}(\cdot) = \mu^{**}(\cdot)$$

This means that the marginal cost of output under a regime of competition in the input market is identical to the marginal cost of output under a regime of perfect price discrimination in the input market. This makes sense because the marginal cost of the input is the same in either case. Note that we have not concerned ourselves with monopoly behavior in the output market. However, by inference, if the marginal cost of output is the same in the two cases, the output decision of the monopolist must be the same.

Does Free Agency Mean Perfect Competition?

Next let's examine the behavior of the owners in the event of free agency. There are different ways that one might view the event of free agency on the behavior of the owners. One is that while they no longer have the power to perfectly monopsonize the players, they still have monopsony power. That is, while the owners cannot extract all labor surplus from the players as they could under the reserve clause, they can extract some surplus. At a minimum, the owners can recognize that the marginal wage of players is increasing as the amount of labor increases. Hence, owners under free agency act as single price monopsonists.

The cost minimizing objective function for a simple monopsonist is:

$$\min C = w_K K + \lambda L^2 + \mu(q - K^\alpha L^\beta)$$

because the simple, single price monopsonist sees the upward sloping supply curve and the cost of labor is this supply function times the amount of labor employed. The demand curve of labor holding output constant is:

$$\tilde{L} = \left(\frac{\tilde{\mu} \cdot \beta \cdot q}{2\lambda} \right)^{1/2}$$

¹⁴ You should derive this result for yourself.

which differs only slightly from the perfect price discriminating monopsonist. The crucial element is the marginal cost function of output. Simple derivation shows that

$$\tilde{\mu} = \hat{\mu} \cdot 2^{\frac{\beta}{2-\beta}}.$$

Thus, the marginal output cost curve of the owners under free agency is higher than the marginal cost under the reserve clause (and higher than it would be under perfect competition). Hence, the profit maximizing output price is higher.

Theory versus Evidence

As an empirical matter, free agency has done much to professional sports. Most observers claim that:

1. Players move more.
2. Players are paid more.
3. Ticket prices are higher.
4. More players are used.

These observations need careful scrutiny to determine their accuracy. However, with sufficient care, we should be able to know the empirical dimension of the change in regime. Given the facts, we are presented with the challenge of crafting a theory to explain them. The theory of single-price monopsony under free agency explains (3) for sure and (2) for the average player. It has nothing to say about (1), and is in direct contradiction to (4).

By contrast, to say that free agency does nothing but replace perfect monopsony with competition in the input market is in contradiction with (4) and (3), says nothing about (1), and is consistent with (2) based on the average player.

These stylized and limited theories do not take into consideration the singular fact that is repeatedly pointed out by league owners: That is, the league is composed of heterogeneous markets. The idea of the reserve clause in its public interest motivation is to spread the talent so that it does not bunch up in the major markets. It is argued that if the talent all goes to the major markets, the total value of the league declines. How this affects the predictions of a model of the sports market is unclear.

¹ ***** **This SAS program examines the equilibrium input and output markets**

***** following an exogenous price increase in one input allowing for

***** a positively sloped supply of another input;

Data flowes;

s1=.1111111111; s2=.3333333333; s3=.5555555556; *shares initialized to original problem;
 rho=-2; sigma=1/(1-rho); *elasticity of substitution;
 eta=-1; a=45; *output elasticity and initial expenditure;
 kount=0;
 do lambda=.1 to 2.5 by .1; *slope of labor supply;
 w_sally=5; w0=w_sally; w_flwr=15; w_ovrhd=25; *initial input costs;

mc=(s1**sigma*w_sally**(-rho*sigma)

```

+s2**sigma*w_flwr**(-rho*sigma)
+s3**sigma*w_ovrhd**(-rho*sigma))
**(1/(-rho*sigma));                                *Marginal Cost;

mktq=a*mc**eta;                                    *Market output;
sallys=1; flwr=1; ovrhd=1;                          *initialize use of inputs;
output;                                             *SAS code to write obs to dataset;
w_flwr=30;                                          *this is the exogenous increase in the cost of cut flowers;
do kount=1 to 100;
w_sally=w0*sallys**lambda;                          *labor supply function. wage rate changes;
                                                    *as input use changes;

mc=(s1**sigma*w_sally**(-rho*sigma)
+s2**sigma*w_flwr**(-rho*sigma)
+s3**sigma*w_ovrhd**(-rho*sigma))
**(1/(-rho*sigma));
mktq=a*mc**eta;
sallys=mktq*(s1*mc/w_sally)**sigma;                *new labor cost affects MC and output;
sally=(s1*mc/w_sally)**sigma;
flwr=mktq*(s2*mc/w_flwr)**sigma;
ovrhd=mktq*(s3*mc/w_ovrhd)**sigma;                *new labor cost also affects input usage;
epsilon_flwr=flwr-1;                               *calculates industry level labor demand elasticity;
output;                                             *Writes observation to dataset;
end;                                                *Iterates to achieve equilibrium in labor and output mkts;
end;                                                *Change labor supply elasticity;

/*
proc gplot; plot w_sally*kount; symbol i=spline;    *These plots check for convergence;
proc gplot; plot mc*kount; symbol i=spline;        *They are suppress with the slash-star;
proc gplot; plot sally*kount; symbol i=spline;     *once convergence is confirmed;
proc gplot; plot mktq*kount; symbol i=spline;
run;*/

data flower; set; if kount=100; invlambda=1/lambda; *Limit d/s to converged values;
sally2flwr=sallys/flwr;
sally2ovrhd=sallys/ovrhd;
proc gplot; plot w_sally*invlambda; symbol i=spline; run; *These plots show equilibria;
proc gplot; plot sally*invlambda ; symbol i=spline; run;
proc gplot; plot sally2flwr*invlambda ; symbol i=spline; run;
proc gplot; plot sally2ovrhd*invlambda ; symbol i=spline; run;
proc gplot; plot mc*invlambda; symbol i=spline; run;
proc gplot; plot mktq*invlambda ; symbol i=spline; run;
proc gplot; plot epsilon_flwr*invlambda; symbol i=spline; run;

2 *****This SAS program depicts long-run supply of output resulting from
*****exogenous shifts in output demand in the case where one input has a
*****positively sloped supply;
*****It follows the program that depicts the effect of an exogenous increase in
*****an input price, so only the code changes are annotated;

Data flowers;
s1=.1111111111; s2=.3333333333; s3=.5555555556;
rho=-2; sigma=1/(1-rho);

```

```

eta=-1; do a=45 to 60 by 1;                                *shifts in output demand;
kount=0;
lambda=2;
w_sally=5; w0=w_sally; w_flwr=15; w_ovrhd=25;
mc=(s1**sigma*w_sally**(-rho*sigma)+s2**sigma*w_flwr**(-rho*sigma)
+s3**sigma*w_ovrhd**(-rho*sigma))**(1/(-rho*sigma));
mktq=a*mc**eta;
sallys=1; flwr=1; ovrhd=1;
output;
do kount=1 to 100;
w_sally=w0*sallys**lambda;
mc=(s1**sigma*w_sally**(-rho*sigma)+s2**sigma*w_flwr**(-rho*sigma)
+s3**sigma*w_ovrhd**(-rho*sigma))**(1/(-rho*sigma));
mktq=a*mc**eta;
sallys=mktq*(s1*mc/w_sally)**sigma;
sally=(s1*mc/w_sally)**sigma;
flwr=mktq*(s2*mc/w_flwr)**sigma;
ovrhd=mktq*(s3*mc/w_ovrhd)**sigma;
epsilon_flwr=flwr-1;
output;
end;
end;

proc gplot; plot mc*kount; symbol i=spline;                *checks for convergence;
proc gplot; plot mc*mktq; symbol i=spline;
run;

data flower; set; if kount=100; invlambda=1/lambda;
sally2flwr=sallys/flwr;
sally2ovrhd=sallys/ovrhd;

proc gplot; plot mc*mktq; symbol i=spline; run;
proc gplot; plot w_sally*sallys ; symbol i=spline; run;

```