

6

The Social Costs of Input-Market Distortions

By DANIEL WISECARVER*

This paper examines the welfare effects of taxes (or other distortions) on factors of production. The analysis will demonstrate two major propositions: 1) Most importantly, given the usual *derived* demand curve for a productive factor, the social cost of a tax on that factor is correctly, completely, and most readily measured by the relevant area between the factor's demand and supply schedules. 2) Under variable proportions, the social cost that is so measured will entail two terms—corresponding precisely to the scale and substitution effects of derived demand theory—both of which represent true deadweight losses.

Although these propositions will be obvious to the practitioner of applied welfare economics, each has nevertheless been the source of considerable confusion in the recent literature. The first confusion is one of double counting,¹ as is explicitly illustrated in the recent paper by Ann Friedlaender (1971) in which the social cost of "value-of-service" (rather than marginal cost) pricing of U.S. rail services is estimated. Friedlaender's calculation consists of two distinct steps. First, the social costs that are measurable in the directly distorted rail markets are determined.

Then, to this direct cost component is added a further element of deadweight loss that is alleged to arise from the fact that all goods requiring rail transport as an input are themselves nonoptimally priced—"indirectly taxed"—due solely to the distortion of rail prices.² We shall show, in Sections I, II, and III, that this sort of summing procedure is unequivocally wrong.

It might, however, be informative to speculate as to the source of this double-counting error. It is conceivable that the confusion has arisen from an eclectic survey of the two outwardly distinct approaches that have in the past been used to determine the social welfare costs of input-market distortions. The first approach, exemplified by the work of Albert Fishlow and Paul David, Jagdish Bhagwati and V. K. Ramaswami, and Harry Johnson, analyzes the problem within the context of the two-sector model of general equilibrium. It uses the concept of a social welfare function to depict the aggregate loss of welfare that is occasioned by an input-market distortion. This is defined to be the decrease in utility that society incurs when it is forced from an indifference curve which is tangent to the

* Assistant professor of economics, Ohio State University. The current version of this paper, completed while I was a staff member of the Office of Tax Analysis, U.S. Department of the Treasury, is a condensation of one chapter of my Ph.D. dissertation at the University of Chicago. I would like to thank and absolve A. C. Harberger, H. G. Lewis, and Larry Sjaastad for their extensive help and comments.

¹ This confusion also appears in Walter Oi and A. P. Hurter, Jr., pp. 340-41, although they do not attempt detailed calculations.

² Friedlaender's direct cost measure is calculated via $L = \frac{1}{2} \sum_i E_i Z_i \delta_i$, where E_i is the elasticity of demand for rail services for commodity type i , Z_i is the proportion of total rail revenues received in i , and δ_i is the distortion of the rail price of i . Similarly, her second cost component is $L = \frac{1}{2} \sum_j v_j \mu_j \gamma_j$, where v_j is the elasticity of demand for commodity j , μ_j is the proportion of the economy's value-added that is generated by industry j , and γ_j is the price distortion for j caused by δ_i , i.e., γ_j is the product of the cost share of rail services in j and δ_i .

economy's undistorted transformation function to a lower social indifference curve which may or may not be tangent to the inferior transformation function that is dictated by the distortion. The second approach, exhibited by Albert Rees and Arnold C. Harberger (1964a, 1966), applies the current reincarnation of the theory of economic surplus directly to the distorted input markets themselves. This method treats the welfare cost of an input distortion as a simple money measure of the deadweight loss that the distortion imposes on the sum of the areas analogous to consumers' and producers' surpluses.

Among the apparent methodological differences between these approaches, one likely to receive considerable attention is the following: on the one hand, the surplus method treats the distortion as causing only one misallocative effect, that being the creation of a divergence between the input's demand and supply prices; on the other hand, the two-sector method explicitly analyzes two separate, distortion-induced effects. In Johnson's words:

Of course, the total economic loss from factor market distortions can only be evaluated by taking into account, in addition to the inefficiency of production, the distortion of consumer choices due to the factor market distortion in producing a divergence of private from social opportunity costs in consumption.

[p. 697]

Thus, even though these alternative approaches are applied to the same (qualitative) problem, our eclectic surveyor might question the compatibility of their results, feeling that the surplus approach captures only part of the total cost that is implied by the two-sector approach. In other words, should not the surplus method additionally allow for the fact that a (positive) distortion in the market for an input will cause the price of any product employing that input also to be set at an artificially high level, thereby imposing

further deadweight losses? The correct answer is the negative. From the viewpoints of both derived demand theory and, at a significantly different level, the "distortion" approach to applied welfare economics, any attempt to add to the input-market measure of social cost a further element of loss that is alleged to occur in the output market when the supply price of the industry's product is forced to an (indirectly) distorted level will entail double counting of the actual costs involved.

That our second major point has been a source of confusion is exhibited by Richard Schmalensee's analysis of the welfare gains accruing from a fall in the price of a factor of production (of course, this is simply the obverse of our analysis of a tax which will raise the price of the factor in question) both within the market for that factor and within the market for the product that employs the factor. He finds, correctly, that under fixed proportions both measures of welfare gain are identical so that *either* measure could properly be used in a social accounting framework. However, when proportions are variable, the input-market measure exceeds that found in the product market by a term which is a pure substitution-in-production effect, as can easily be seen by suitable manipulation of Schmalensee's equation (7). From here, Schmalensee concludes that only the product-market measure should be used, contending that the factor-market measure will overstate the gain in welfare that is due to the decrease in factor price. Implicitly, therefore, he concludes that the "substitution" component of the factor-market measure does not represent true social gain. Sections II and III of the present paper will show that Schmalensee's conclusion is incorrect.³

³ The recent, widely discussed study by Stephen P. Magee, p. 658, in which Schmalensee's faulty conclusion

In the analysis to follow, Section I refutes the double-counting error in terms of derived demand theory, utilizing the case of fixed proportions for ease of diagrammatical illustration. Section II, in treating the more complicated case of variable proportions, not only reaffirms our conclusion as to double counting but also shows that the substitution component of the factor-market measure of social cost represents a true element of deadweight loss. Finally, in Section III we review the entire analysis from the more intuitive standpoint of applied welfare economics.

I

The whole of the analysis to follow is based upon the standard, received theory of derived demand; factors of production do not, in and of themselves, yield utility to their demanders. Rather, inputs are demanded solely⁴ because they can be used to provide those finished products that do yield utility to their purchasers. Thus, any utility that might be attributed to the area under a factor's demand curve is necessarily a mere reflection of the utility that its services ultimately provide. The word reflection here is very important, in that the utility garnered by the ultimate purchasers of the product of a factor's services is reflected in the demand curve for that factor. Obviously we do not claim that the total utility to be gained from the production of a good is the sum of the direct gain, as measured under that good's demand curve, and the reflected gain, as measured under the demand curve

is favorably cited, gives further indication that this second source of confusion warrants clarification.

⁴ It might be objected that certain inputs, particularly intermediate goods, are demanded in the aggregate partly for their productive services but also partly for final consumption purposes. However, since the major point of this paper is that the best place for calculating the social costs of input taxes is the affected input markets themselves, this sort of objection can only strengthen our argument.

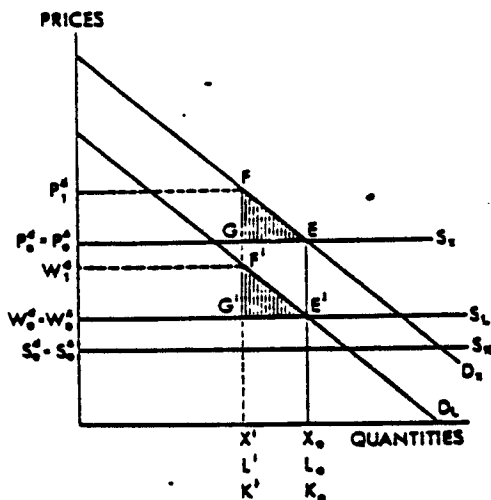


FIGURE 1

for factors used to produce the good. But then, by applying this obvious proposition in reverse, we have established a major point of this paper: it is incorrect to calculate utility losses in both the input and the output markets when only the input is explicitly taxed.

This conclusion is most easily seen when production technology is characterized by fixed proportions, an instructive, if not entirely realistic, framework of analysis. Hence, in the discussion to follow, we define units such that one unit of output (X) requires one unit each of labor (L) and capital (K), i.e., $X = \min(L, K)$. Consider Figure 1. The derived demand curve for L is constructed as the vertical difference between the demand for X and the supply of K , here assumed to be perfectly elastic as is the supply of L ; the supply schedule for X is the vertical sum of the supplies of L and K . Overall equilibrium occurs where the demand and supply prices of X , L , and K equal each other, respectively, and where $P^d = w^d + s^d$ and $P^d = w^d + s^d$.⁵ Thus

⁵ In the notation of this section, P , w , and s are the prices of X , L , and K , respectively, and the superscripts d and s denote demand and supply prices. For a more

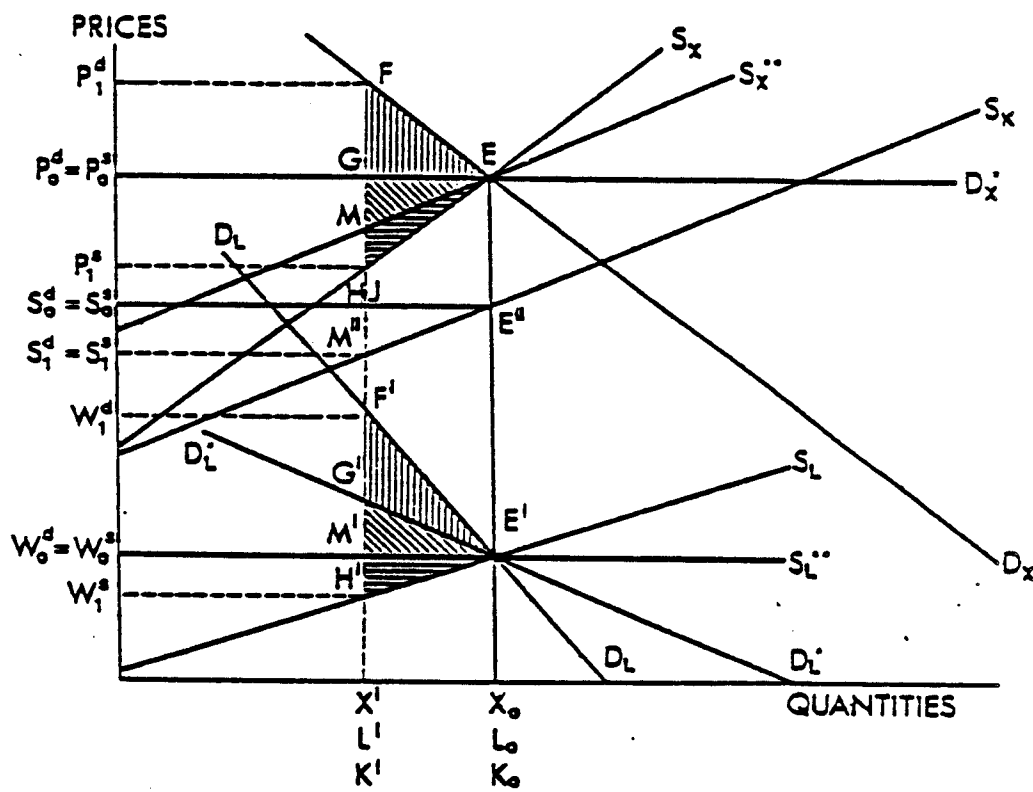


FIGURE 2

the points E and E' represent initial, undistorted equilibrium.

Suppose now that a tax of $F'G' = T$ per unit L is imposed, forcing equilibrium to shift to point F where $P_1^d = w_1^d + s_1^d = w_0^d + T + s_0^d$ and $P_1^s = w_1^s + s_1^s = P_0^s$. By the usual calculation of deadweight losses, the social cost of T in the market for L is $\frac{1}{2}T(L_0 - L_1) = E'F'G'$. If, however, one were to seek in the market for X a measure of social cost that might be attributed to the indirect distortion of P due to T , that measure would be $\frac{1}{2}(P_1^d - P_0^d)(X_0 - X_1) = -\frac{1}{2}T\Delta X = EFG = E'F'G'$. In the case depicted in Figure 1, where we assume that there are no rents to be lost, the dead-

weight cost of a tax can only be in terms of consumption utility—consumers' surplus—and this cost is equally well measured by EFG or $E'F'G'$. But to assert that the total deadweight cost is the sum of EFG and $E'F'G'$ is to engage in exact double counting.

This property of exact double counting also holds when rents to either or both factors are involved, as will be seen with reference to Figure 2. As before, D_L is constructed as the vertical difference between D_x and S_x , and S_x is the vertical sum of S_L and S_x'' (momentarily ignore the curves D_x' , D_L' , S_L'' , and S_x''). Initial equilibrium at E is determined by the intersection of S_L and D_L at E' . If a tax of $F'H' = T$ is now imposed on labor, equilibrium shifts from E to F (E' to F') and

detailed description of the geometrical tool being used in this section, see Milton Friedman, ch. 7.

the resultant social cost is of course $-\frac{1}{2}T\Delta L = E'F'H'$. Again, due to the additive properties of demand and supply prices under fixed input proportions, $E'F'H'$ must equal the cost of T as it might be measured in the X market, i.e., must equal $-\frac{1}{2}(P_1^d - P_1^s)\Delta Y = EFH$. For at F we have

$$P_1^d = w_1^d + s_1^d = w_1^s + T + s_1^s$$

and

$$P_1^s = w_1^s + s_1^s$$

so that

$$P_1^d - P_1^s = T.$$

In order to show that the labor-market measure of social cost ($E'F'H'$) neglects no element that would require the inclusion of (any part of) the product-market measure (EFH), we shall break these triangles down to their component parts and see to what each component corresponds. To this end, it is convenient to consider the following hypothetical experiment. Suppose that the goal of the taxing authority is not the collection of revenue but is rather, for whatever reason, the restriction of the use of labor in X to L' , thereby allowing T to vary inversely with the elasticities of S_L and D_L . Thus, given D_X , S_K , and S_L , the tax set is $F'H'$. Alternatively, if the demand for X were D_X^0 , the derived demand for L would be D_L^0 . If in addition the supply of L were S_L^0 (implying a supply of X of S_X^0), then the tax would be set at $G'M'$ per unit labor. Now with these assumptions, there is obviously no consumers' surplus and no rent to labor to be lost, regardless of where we choose to measure social costs. However, there still is a dead-weight cost of $E'G'M'$, which is easily shown to equal the areas EGM —producers' surplus as seen in the X market—and $E''JM''$, the rent, as seen in the K

market, that is lost by owners of capital and yet is not transferred to any other economic entities. Thus, the area $E'G'M'$ in the directly taxed labor market captures the deadweight loss of rent that is imposed upon the untaxed factor.

If now the supply of labor is taken to be S_L , the authority will increase the tax on L to $G'H'$, which increase will entail no change in the position of capital owners. It does, however, increase the welfare cost by $E'M'H'$ at the expense of owners of labor. It is also clear that, with D_X^0 and S_X , $P_0^s - P_1^s = G'H'$, so that $EGH = E'G'H'$ and therefore that $EMH = E'M'H'$.

Finally, let demand for X return to D_X and the tax to $F'H'$. This change does not further affect factor owners, but it does lead to a loss in consumers' surplus of EFG . And, since we have already shown that $EFH = E'F'H'$ and that $EGH = E'G'H'$, it follows that $EFG = E'F'G'$.

To summarize, the social cost of a tax on a factor of production in the case of fixed proportions (and when there are no other distortions in the economy) can be measured either in the directly distorted factor market or in the indirectly distorted output market; the market so chosen is a matter of indifference. Furthermore, we can divide the input-market measure into areas that are analogous to the traditional division of welfare cost into producers' and consumers' surpluses. The area above the taxed input's supply schedule and below its initial price—the analogue to producers' surplus—is rent lost by that input which is not gained by other sectors of the economy. The analogue of consumers' surplus, on the other hand, contains both the net social loss in rent to the untaxed factor⁶ and the consumption utility lost.

⁶ With fixed proportions, the argument can obviously be extended to n inputs. Then, the social loss of rent (corresponding to $E''JM''$ in Figure 2) to all of the $n-1$ untaxed factors will be counted in the consumers' surplus analogue below the taxed factor's demand curve (corresponding to $E'G'M'$ in Figure 2).

because of the change in the price of the final good that is caused by the taxation of one of its inputs. The great error to be avoided is to measure tax-induced deadweight losses in the factor market and then add to these losses any element of further social cost that is alleged to stem from the product market.

II

Although the case of variable proportions is more complicated, at least from the purely mathematical viewpoint, our basic conclusion with respect to the error of double counting still holds. Furthermore, the analysis is enriched because of the dual nature of the welfare effects of factor taxes; i.e., factor taxes entail both a loss of "reflected" utility due to the consequent product-price distortion and a waste of potential output due to the socially nonoptimal input mix that the tax will induce profit-maximizing firms to utilize.

Assume that in a competitive industry (composed of π identical firms) we have the linearly homogeneous production function⁷

$$(1) \quad X = F(L, K)$$

and that w and r , the prices of L and K , are fixed to the industry. Then it follows that⁸

$$(2) \quad \eta_L = \alpha_L \eta - \alpha_K \sigma$$

where η_L is the elasticity of derived demand for labor, the α terms represent the cost shares of the respective factors, η (a negative number) is the elasticity of demand for X , and σ (a positive number) is the elasticity of substitution between L and K . If, from a point of initial equilib-

⁷ The same results obtain for any production function if the industry is in a position of long-run equilibrium. That is, in long-run equilibrium competitive firms operate at the minimum point of their average cost curves, and here the relevant production parameters are (locally) the same as those of constant returns to scale.

⁸ See, for example, R. G. D. Allen and J. R. Hicks.

rium, L is taxed at the percentage rate t , it follows that the resultant welfare-cost triangle is measured by

$$(3) \quad WCL = -\frac{1}{2}w_0L_0\eta_L t^2 \\ = -\frac{1}{2}w_0L_0\alpha_L\eta^2 t^2 + \frac{1}{2}w_0L_0\alpha_K\sigma t^2$$

Notice that these two deadweight terms correspond exactly to the scale (reflected utility) and substitution effects.

Now in order to determine the indirect social cost of t that might be measured in the market for X , we must know the effect of t on the supply schedule of X . By the condition of competitive equilibrium that price equal marginal and average cost, and by the properties of the assumed production function, it follows that

$$(4) \quad EP_x = \alpha_L Ew + \alpha_K Er$$

where E is the operator d/d . Fixed factor (supply) prices gives us $Er=0$ and $Ew=t$, so that

$$(5) \quad EP_x = t_{xL} = \alpha_L t$$

where t_{xL} represents the indirect or implicit tax on X that is due to t . Then the implicit welfare cost of t that is brought about by the labor-tax-induced shift of the industry's supply schedule for X is

$$(6) \quad WCs = -\frac{1}{2}P_{xx}X_0\alpha_L^2 t^2 = -\frac{1}{2}w_0L_0\alpha_L^2 t^2$$

From (3) and (6), it follows that

$$(7) \quad WCL = WCs + \frac{1}{2}w_0L_0\alpha_K\sigma t^2$$

Equation (7) shows us two important points. First, when factors can be substituted for each other (i.e., when $\sigma > 0$), WCL exceeds WCs by a term that is a pure "substitution-in-production" effect. It is this term that would be omitted if the social cost of t were calculated, as, for example, Schmalensee has suggested, solely in the output market. Indeed, in the extreme case of a zero elasticity of demand for X , a traditional "triangle-analysis" in the output market would lead to the incor-

rect conclusion that the social cost of a tax on labor employed in the X industry is zero. The fact of substitution implies that, even if there is no scale effect in the demand for labor, a tax on labor will induce an inefficient movement to a more capital-intensive production of X ; the resultant inefficiency cost is captured automatically when the analysis is carried out within the taxed labor market itself.

The second point emerging from equation (7) is that the labor-market measure of the tax's social cost also captures the "scale effect" when there is a scale effect, i.e., when $\eta \neq 0$. That is, WC_s is included in WC_L . Therefore, any attempt to add to WC_L the deadweight loss that one might conceive as occurring in the output market will entail an exact double counting of WC_s . So long as the relevant demand curve for labor is the usual derived demand, the welfare cost of a tax on labor is completely measured solely within the labor market.

There is no difficulty in interpreting the scale effect (WC_s) of WC_L ; WC_s is simply the reflection of the deadweight loss of consumers' and producers' surpluses that is imposed on the economic agents in the X market when a tax on labor forces P_x to an artificially high level. However, the fact that the "substitution component" of WC_L — $\frac{1}{2}w_0L_0^2\alpha_K\sigma$ —is also a genuine element of social cost apparently requires further elaboration. There are at least three methods of explaining the mechanism that imposes this extra deadweight loss upon society. First, and easiest, we could simply appeal to the two-sector model of general equilibrium and assert that the substitution component of WC_L captures the essence of the inefficiency-in-production effect, while WC_s captures the distortion-of-consumer-choices effect (see the quotation from Johnson, above). However, even though this explanation has a great deal of intuitive appeal, it can-

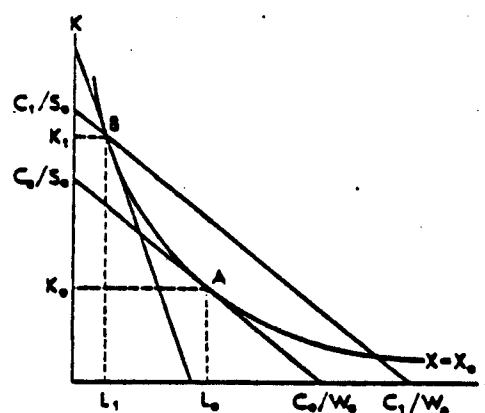


FIGURE 3

not withstand rigorous examination on its own merit, since our measure of WC_L was not explicitly derived from the general equilibrium model. Fortunately, the other two methods of explanation allow us to avoid any such "proof by assertion."

The second method relies on the fact that when technology allows input substitution, a tax placed on just one factor of production will lead profit-maximizing entrepreneurs to utilize an input mix which does not minimize social factor costs. For example, in Figure 3 we have the industry's isoquant map. Prior to the tax on labor, industry output X_0 was produced with the socially efficient input mix indicated by point A at total cost $C_0 = w_0L_0 + s_0K_0$. After the tax, X_0 would be produced with the input mix indicated by point B, implying a total industry cost of $w_0(1+t)L_1 + s_0K_1$. However, at point B private and social factor costs diverge. Since we are still assuming infinitely elastic factor-supply schedules, the factors' supply prices, and therefore their true social costs per unit, remain at w_0 and s_0 ; nevertheless, the tax has caused the industry to become more capital intensive than is socially optimal given w_0 and s_0 . Therefore, the total social cost—total private

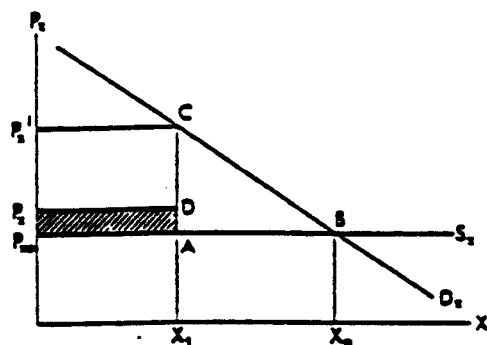


FIGURE 4

costs less tax payments—of X_0 when produced at input mix B exceeds C_0 ; i.e.,

$$C_1 - C_0 = w_0 L_1 + r_0 K_1 - w_0 L_0 - r_0 K_0$$

It can be shown that when $C_1 - C_0$ is approximated up to the quadratic terms of a Taylor series expansion about $t=0$, the resultant excess total factor cost is the substitution component of WC_L which is given in equation (7).

The key to understanding the third method—which is essentially a corollary of the preceding explanation—is to compare the total tax revenue that could be raised under two alternative taxation policies: (a) a tax on labor of t percent, and (b) a tax on X of $\alpha_L t = t_{XL}$ percent. That is, in this exercise we transform the tax t on labor into an implicit tax t_{XL} on X and compare the implied revenue (in terms of X) with the revenue (also in terms of X) that would be garnered if X were taxed directly at the percentage rate t_{XL} . The excess of revenue implied by policy (b) over that implied by policy (a) turns out to be $\frac{1}{2} w_0 L_0^2 \alpha_L t^2$, which is again the substitution component of WC_L .

The analysis is depicted in Figure 4. A direct tax on X of t_{XL} percent will raise the price facing consumers from P_{00} to P_1' and will generate revenue of $P_{00} P_1' CA$.⁹

⁹ Notice that under fixed proportions the revenue $P_{00} P_1' CA$ is exactly equal to the revenue that would be

On the other hand, a tax of t percent levied on labor will cause the price of X again to rise from P_{00} to P_1' ; however, the factor tax alters the relative prices of labor and capital facing the industry. As can be seen by reference to Figure 3,¹⁰ when the industry moves to more capital-intensive means of production the after-tax (but net-of-tax-payments) supply price of X exceeds P_{00} at every level of output. Therefore, part of the area $P_{00} P_1' CA$ which was tax revenue when X was directly taxed will become factor payments when labor is taxed instead, leaving the smaller area $P_1' P_2' CD$ available as tax revenue. In other words, both tax policies induce the industry to increase the price at which it is willing to supply X from P_{00} to P_1' . When X is taxed directly, all of the price increase, for each unit sold, becomes tax revenue. But when labor is taxed, part of the price increase reflects the higher supply price of X , at social factor cost, that results from the socially nonoptimal combination of inputs that the industry is induced to employ. Therefore, only part of $(P_1' - P_{00})$ actually becomes tax revenue, and the area $P_{00} AD P_1'$ is the deadweight cost of producing X_1 at the socially nonoptimal input mix that results from taxing labor. Algebraically, this cost can be expressed as

$$\begin{aligned} P_{00} AD P_1' &= (P_1' - P_{00}) X_1 \\ &= (P_1' - P_{00}) X_1 - w_0 L_1 t \end{aligned}$$

Expansion of the latter expression to the

garnered from the tax t on labor, as measured in the labor market. These same qualitative results, i.e., that the welfare costs and revenue collections of t and t_{XL} (levied directly on X) are identical, obtains regardless of σ if $\sigma_X = 0$. See the discussion of equation (16) below.

¹⁰ The translation from Figure 3 to Figure 4 follows once it is recognized that for each level of output, i.e., for each isoquant, the tax on labor implies an excess social as well as private cost of production. And, since our analysis is one of comparative statics, price equals average cost both before and after the tax; P_1' is the tax-inclusive average cost, P_2' the tax-exclusive average cost.

quac
our;
H
poss
of t'
(pos
for
ligh:
ulat
subs
ket
mea
nati
welf
reas
whe
not
simp
tort
mar
all c
but
in a
T
com
just
tion
whe
ass:
ties
gen
ent
foll
sta.
equ
ma
ten
(8)
(9)
Als
ela
so
"e
eic
eial
tur
reie

quadratic terms of a Taylor series yields our previously determined result.

However, in spite of the fact that it is possible to represent all of WCL in terms of the demand curve, supply curve, and (post-tax) cost conditions in the market for X , it should be evident, especially in light of the requisite mathematical manipulations, that practical measurement of the substitution component in the output market is intractable, at best. Furthermore, measures such as ADP/P_s do not fit naturally into the framework of applied welfare economics since, among other reasons, this component of cost will exist whether the level of output changes or not (see Section III). Thus we see that the simple concept of the social cost of a distorted input market, as measured in the market for that input, not only captures all of the relevant deadweight components but also is the only concept that does so in a reasonably applicable manner.

The discussion so far has ignored any complications of factor rents. However, just as was the case under fixed proportions, none of our conclusions are altered when rents are included, i.e., when we assume noninfinite, factor-supply elasticities. Hence, only an outline of this more general case will be presented here. The entire analysis can be generated from the following four-equation model. Under constant returns to scale (or in long-run equilibrium), a competitive industry's demands for labor and capital may be written as

$$(8) \quad EL = \eta EP_s - \alpha_K \sigma E(w/s)$$

$$(9) \quad EK = \eta EP_s + \alpha_L \sigma E(w/s)$$

Also, we define ϵ_L and ϵ_K as the supply elasticities of L and K facing industry X , so that¹¹

$$(10) \quad EL = \epsilon_L Ew$$

$$(11) \quad EK = \epsilon_K Es$$

By combining (9) and (11) to determine Es (having relied upon (4)) and substituting the result into (8), we can solve explicitly for the elasticity of derived demand for labor,

$$\eta_L = \frac{EL}{Ew} = \frac{\sigma \eta - \alpha_K \epsilon_K \sigma + \alpha_L \epsilon_K \eta}{\epsilon_K + \alpha_L \sigma - \alpha_K \eta}$$

Similarly, this simple factor-markets model, along with (4), gives us an explicit expression for the elasticity of product supply:¹²

$$\epsilon_s = \frac{\alpha_L \epsilon_K + \sigma(\alpha_L \epsilon_L + \alpha_K \epsilon_K)}{\sigma + \alpha_L \epsilon_K + \alpha_K \epsilon_L}$$

Given η_L and ϵ_L , a rate of tax t on labor of course implies a welfare cost of

$$(12) \quad WCL = -\frac{1}{2} w_0 L_0 t^2 \left(\frac{\epsilon_L \eta_L}{\alpha_L - \eta_L} \right)$$

Once again, to determine any implicit deadweight cost that occurs in the product market due to the imposition of t on L , we must calculate $EP_s = t_{sL}$ via (4). However, unlike the earlier case of equation (5), when t induces the industry to shift to more capital-intensive means of production, Es will not be zero, Ew will not simply be t , and t_{sL} will not in general be $\alpha_L t$. Consider Figure 5. Initial production of X_0 is accomplished with L_0 of labor and K_0 of capital when (w_0/s_0) represents relative factor prices (panel A). The tax on labor causes a new factor-price ratio, say $(w/s)'$. Now X_0 could still be produced at the factor combination A , implying a cost increase of $w_0 L_0 t$. Such a situation would emerge (in panel B) as a shift in the supply of X by $\alpha_L t$ percent (from S to S'). However, point A is no longer a cost-minimizing position for the industry; the tax will

¹¹ It should be noted that throughout we are assuming either that we are operating entirely on constant-elasticity curves or that the relevant parametric disturbances are small enough to make any changes in the relevant elasticities negligible.

¹² Both derivations here are almost identical to those found in Richard Muth.

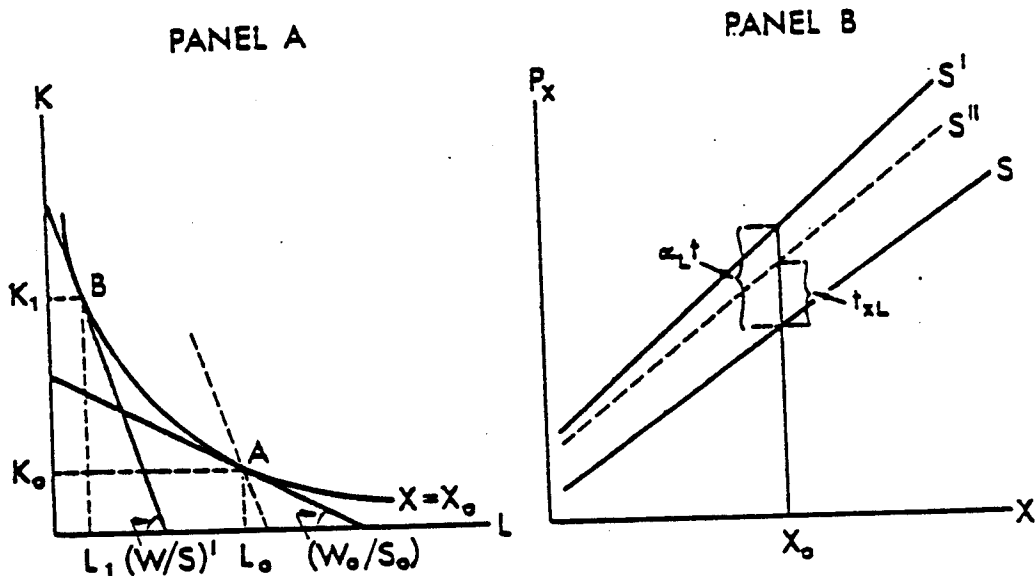


FIGURE 5

induce the new factor combination L_1 and K_1 at point B. Total industry costs will therefore not rise by $\alpha_L t$ percent but by some amount qualitatively represented by t_{xL} in panel B.

The simplest way to find the relevant percentage increase in industry supply price that is occasioned by t is to derive t_{xL} from output-constant factor-demand curves. This procedure has the virtue of yielding an expression that is independent of the elasticity of demand for X . Thus, in equations (8) and (9) we set $\eta EP_x = EY = \text{zero}$. The resultant expressions, along with (11), give us the output-constant elasticity of demand for labor.

$$\eta_L^s = \frac{-\alpha_L \epsilon_L \sigma}{\epsilon_K + \alpha_L \sigma}$$

(here the superscript s denotes "substitution effect only"). In using (4) to define t_{xL} , we must recognize that P_x is the industry supply price, while the relevant w and s are the demand prices that the industry must pay for its inputs. When labor is taxed, the demand and supply

prices of labor (w^d and w^s) differ by the absolute amount of the tax; therefore, to derive t_{xL} we must use the percentage change of w^d . Starting from a no-tax situation, it is easily shown that

$$(13) \quad E_{w^d} \Big|_{x_0} = \left(\frac{\epsilon_L}{\epsilon_L - \eta_L} \right) t$$

Since there is no tax on capital, its demand and supply prices will always be equal; hence, $E_{s|X}$ is simply derived from (11) and the output-constant version of (9). This result and (13) give us

$$(14) \quad t_{xL} = EP_x = \alpha_L t \cdot \left(\frac{\epsilon_L (\epsilon_K + \sigma)}{\epsilon_L \epsilon_K + \sigma (\alpha_L \epsilon_L + \alpha_K \epsilon_K)} \right)$$

Notice that $t_{xL} = \alpha_L t$, the case that always obtains with fixed factor prices, only when $\epsilon_L = \epsilon_K > 0$.¹²

¹² Notice also that $t_{xL} > \alpha_L t$ when $\epsilon_L > \epsilon_K$ and that $t_{xL} < \alpha_L t$ when $\epsilon_L < \epsilon_K$. While it might seem paradoxical to find that, when substitution is possible, the supply-price increase can actually exceed that which is possible when proportions are fixed, the puzzle is removed once we recognize that the individual firms within our com-

Thus, given t , ϵ , and η , the implicit WC_L is

$$(15) \quad WC_L = -\frac{1}{2} P_{xx} X \theta_{xL}^2 \left[\frac{\epsilon + \eta}{\epsilon - \eta} \right]$$

Once (12) and (15) are expanded (and the denominator of WC_L is changed), a comparison reveals that

$$(16) \quad WC_L = WC_L + \frac{1}{2} w_0 L \epsilon^2 \left(\frac{\alpha_{K^2} \sigma_{LK}}{\epsilon_{LK} + \sigma(\alpha_{LL} + \alpha_{KK})} \right)$$

Thus, (16) again shows that all of the deadweight loss of welfare that one might wish to measure in the market for X , due to its artificially high price, is captured in the directly taxed labor market when the labor-demand curve is defined in the usual manner, i.e., as a curve consisting of scale and substitution effects.¹⁴ Furthermore, the substitution component of WC_L , which vanishes when $\sigma = 0$, is independent of η and therefore of X . In addition, notice

petitive industry act as though they faced fixed factor prices. Therefore, when each of them demands more capital relative to labor, the outcome is to shift the industry's demand curve along a positively sloped supply of capital curve. The resultant increase in the price of capital will of course apply not only to the marginal units employed, but to all inframarginal units as well. For further elaboration of this line of analysis, see Friedman, ch. 5.

¹⁴ With variable proportions, extension of the argument to the n -input case must as yet rely on the intuition of derived demand. Regardless of the number of inputs, it is simply incorrect to ascribe to a tax on an input utility losses in both the input and output markets. There is, in addition, some corroborating "circumstantial evidence." Following the model presented by W. E. Diewert, we can develop expressions for the elasticities of derived demand for factor n and for the supply of product X . As should be expected, ϵ_x contains terms in the elasticities of supply of all n inputs and in the partial elasticities of substitution between them. Also, η_n contains terms in the elasticities of supply of the $n-1$ other inputs, their partial elasticities of substitution, and the elasticity of demand for X . The circumstantial evidence is that these expressions (ϵ_x and η_n) reduce to equations that are presented in the text above for $n=2$. Unfortunately, even for $n=3$, the algebra required to carry out the same sort of analysis that would result in an expression analogous to (16) has proved to be prohibitive. (But see fn. 17.)

that when the untaxed factor (capital in the present exercise) is supplied to the industry with zero elasticity, the substitution component vanishes regardless of σ . This is simply a reflection of the fact that, for each level of output, there is only one combination of inputs that will be utilized when the amount of capital available is fixed.¹⁵ Finally, all three of the arguments made above to verify that the substitution component of equation (7) is a genuine element of social cost that is measurable in the market for X also apply to the substitution component of (16).

III

Although the algebraic exercises of the previous sections clearly substantiate our two basic propositions, it is instructive to reconsider the entire problem from the more intuitive standpoint of applied welfare economics. To this end, we rely upon the following general formula for the welfare change that is associated with any policy variable z :

$$(17) \quad WC = \int_{-z}^z \sum_i D_i(z) \frac{\partial X_i}{\partial z} dz$$

where D_i is the absolute size of the distortion between demand and supply prices for activity i , and $(\partial X_i / \partial z) dz$ shows the change in the level of activity i as z is varied.¹⁶ In our present application of this expression, z is the factor tax and all other D_i are assumed to be zero.

But this last assumption immediately proves our case with respect to the double-counting error. The only direct distortion is that in the factor market; while it is of course true that the factor tax causes a shift of the product's supply curve, this

¹⁵ This statement assumes that capital and labor are fully employed both before and after tax. Since any fiscal tool that led directly to unemployment would presumably be excluded from consideration, the assumption appears to be reasonable.

¹⁶ A more detailed exposition, of course, appears in Harberger (1971).

shift occurs in a market which is not itself distorted and hence entails no additional component of welfare cost.¹⁷

Furthermore, the whole of the measure summarized by equation (7) or (16) does in fact represent welfare change. All that expression (17) requires for its validity (at least in terms of our present discussion) is that the distortion represent the difference between freely determined demand price and true (social) marginal cost, i.e., supply price; the characteristics of the relevant activity that underlie and determine demand and supply prices are of no special significance. In particular, the fact that a factor's demand price reflects the possibility of substitution with other inputs in production as well as the demand price for final goods does not in any sense indicate, despite Schmalensee's contention, that some portion of this demand price—or some portion of the area under the factor's demand-price schedule—does not reflect true social value. Therefore, our second basic proposition that the substitution components of equations (7) and (16) are true elements of deadweight cost is shown to be valid as well.

In spite of all the above admonitions against attempting to account for the welfare effects of a factor tax in both the product and factor markets, there is one case in which such a procedure is legitimate. However, its illustration which follows also demonstrates perforce the pitfalls that await any such analysis.

Whenever equation (17) is applied to a distorted system, the effect of each ΔX_i (i.e., of each $(\partial X_i / \partial z) dz$) must be counted only one time. While this is usually assured by considering only *direct* distortions as being relevant, it can also be accomplished, if one insists on treating

both t and t_{22} (to continue the terminology of the previous section) as relevant distortions by removing the scale effect from the demand for labor. That is, this procedure requires that the deadweight cost of t be measured with reference to a substitution-effect-only demand for labor curve. Then, the total welfare cost of t would be correctly measured as, say, WFC , plus some WFC_2 which contained only a term in the elasticity of substitution, as, for example, the second term in equation (7) or (16).

However, we should notice that the practical application of even this "correct" separate accounting of both the explicit tax on labor and the implicit, derived tax on output would require a special degree of caution. If t ultimately results in a social cost of WFC , plus WFC_2 , the ΔX that we would observe in the product market is the proper term for the determination of WFC . However, the ΔL that we would presumably observe in the labor market is not the proper term for WFC_2 . The observed ΔL will be the sum of the change in labor demanded at constant output, due to its tax-induced increase in price, and the change in labor employed that results from the shift of the constant-output demand curve for labor which will necessarily follow ΔX . Now since this latter demand shift occurs in a positively distorted market, it entails a "rectangle" of additional welfare cost, and this rectangle will be translated into a cost component that is exactly WFC_2 .

In order to see this last point, suppose that a given tax on labor is to be put into effect in miniscule steps such that in successive periods the existing tax is, for example, .01 t , .02 t ,99 t , t . Then, after each piece of tax has been imposed, there will result demand and supply prices for labor that diverge by the then-existing, total amount of the tax. The locus of demand prices generated in this manner will

¹⁷ Notice that the reasoning in this paragraph is completely general, regardless of the number of inputs; therefore we need not worry about the algebraic difficulties encountered in fn. 14.

reflect both the substitution away from labor, at each level of output, and the scale effect on output that will result from every successive addition to the demand price of labor. That is, this locus of labor's demand prices will in fact trace out the derived demand curve for labor. Thus, using the observed ΔL for the calculation of WC_2^* is equivalent to using labor's derived demand curve; and as we have seen, this procedure will, in conjunction with considering t_{KL} to be a relevant, separate distortion, overstate the social cost of t by precisely IVC_2 , despite initial efforts to avoid this pitfall of double counting.

One further point should be made in closing. Although our analysis has been explicitly concerned only with the effects of just one distortion in one productive sector of the economy, our results will be fully compatible with models of general equilibrium and with any sort of generally-distorted system. Because we have confined ourselves to the effects of a tax on the use of one input in only one activity, it is clear that the conclusion—i.e., that the relevant measure of welfare cost is completely captured by the deadweight-loss triangle below the derived demand curve in the input market itself—holds regardless of what that activity and/or input might be. It is therefore but a small step to conclude that our results must hold when we consider the welfare effects of a system of taxes that imposes differential rates upon different factor uses or upon different factors employed in the same activity, or of a system which contains both sorts of discriminatory elements.

Once we recognize the argument's generality, our analysis of the social costs of such a general system of distortions need proceed no farther than Harberger's model (1964a); for in that model Harberger explicitly recognizes that his "reduced form," reaction coefficients— $R_{ij} = \partial X_i / \partial T_j$, where the X_i represent the number of units of

factor X used in activity i —"... reflect not only conditions of final demand and supply, but also conditions of factor substitution" (p. 68). We might also add that there is no reason for restricting ourselves to just two inputs which can be substituted for each other in production.¹³ It has been the aim of this paper to clarify more explicitly the rationale behind the proper framework for analyzing the welfare effects of input-market distortions.

¹³ By assuming that all inputs other than capital and labor are used in fixed proportions to their respective products, Harberger can express the social cost of an entire system of product and input taxes as $\Delta W = \frac{1}{2} (\sum_i \sum_j G_{ij} B_i B_j + \sum_i \sum_j H_{ij} B_i E_j + \sum_i \sum_j M_{ij} E_i E_j + \sum_i \sum_j N_{ij} F_i F_j)$, where B and E are taxes on the various uses of capital and labor, respectively, and $G_{ij} = \partial K_i / \partial B_j$, $H_{ij} = \partial K_i / \partial E_j$, $M_{ij} = \partial L_i / \partial E_j$ and $N_{ij} = \partial L_i / \partial F_j$. If we wish to allow for more than two inputs whose (partial) elasticities of substitution with each of several other inputs is nonzero, we can generalize Harberger's expression to $\Delta W = \frac{1}{2} \sum_i \sum_j \sum_k F_{ijk} V_j V_k$, where there are m variable inputs under consideration. Here, $F_{ijk} = \partial F_i / \partial V_j \partial V_k$ represents the change in the i th activity's use of input k due to a tax on input j in activity i . These changes are to be understood to allow for the rest of the sectors in the economy to adjust to their new equilibrium levels. For $m=2$ (K and L), ΔW reduces to ΔW .

REFERENCES

- R. G. D. Allen, *Mathematical Analysis for Economists*, New York 1938.
- J. Bhagwati and V. K. Ramaswami, "Domestic Distortions, Tariffs, and the Theory of the Optimum Subsidy," *J. Polit. Econ.*, Feb. 1963, 71, 44-50.
- W. E. Diewert, "A Note on the Elasticity of Derived Demand in the N-Factor Case," *Economica*, May 1971, 38, 192-98.
- A. Fishlow and P. A. David, "Optimal Resource Allocation in an Imperfect Market Setting," *J. Polit. Econ.*, Dec. 1961, 69, 529-44.
- A. F. Friedlaender, *The Dilemma of Freight Transport Regulation*, Washington 1969.
- , "The Social Costs of Regulating the Railroads," *Amer. Econ. Rev. Proc.*, May 1971, 61, 226-34.
- M. Friedman, *Price Theory: A Provisional Text*, Chicago 1962.
- A. C. Harberger, (1964a) "The Measurement

- of Waste," *Amer. Econ. Rev. Proc.*, May 1964, 54, 63-76.
- , (1964b) "Taxation, Resource Allocation and Welfare," in *The Role of Direct and Indirect Taxes in the Federal Revenue System*, Nat. Bur. Econ. Res. and Brookings Inst. conference report, Princeton 1964.
- , "Efficiency Effects of Taxes on Income from Capital," in M. Krzyaniak, ed., *Effects of the Corporation Income Tax: Papers Presented at the Symposium on Business Taxation*, Detroit 1966.
- , "Three Basic Postulates for Applied Welfare Economics," *J. Econ. Lit.*, Sept. 1971, 9, 785-97.
- J. R. Hicks, *The Theory of Wages*, 2d ed., New York 1963.
- H. G. Johnson, "Factor Market Distortions and the Shape of the Transformation Curve," *Econometrica*, July 1966, 34, 686-98.
- S. P. Magee, "The Welfare Effects of Restrictions on U.S. Trade," *Brookings Papers*, Washington 1972, 3, 645-701.
- R. F. Muth, "The Derived Demand Curve for a Productive Factor and the Industry Supply Curve," *Oxford Econ. Pap.*, July 1964, 16, 221-34.
- W. Y. Oi and A. P. Hurter, Jr., *Economics of Private Truck Transportation*, Dubuque 1965.
- A. Rees, "The Effects of Unions on Resource Allocation," *J. Polit. Econ.*, Oct. 1963, 71, 69-78.
- R. Schmalensee, "Consumer's Surplus and Producer's Goods," *Amer. Econ. Rev.*, Sept. 1971, 61, 682-87.