Monopolistic Two-Part Pricing Arrangements

Richard Schmalensee

WP# 1105-80 Revised: October 1980

Monopolistic Two-Part Pricing Arrangements

Richard Schmalensee*

This essay is mainly concerned with Walter Oi's Disneyland problem: pricing a fixed input (admission to the park, or Polaroid cameras) and a variable input (individual rides, or Polaroid film) to maximize profit, though profit-constrained welfare maximization is also treated. The structure of demand in such situations is fully described when customers are either households or competitive firms. The implications of customer diversity and other market attributes for optimal policies are presented. The welfare properties of single-price and two-part tariff monopoly equilibria are compared, and potential welfare gains from tying contracts are discussed.

*Sloan School of Management, Massachusetts Institute of Technology

I am indebted for comments on an earlier version of this paper to Thomas Stoker, an anonymous referee, seminar audiences at MIT, the Universities of Pennsylvania and Florida, and Dartmouth College, and participants in a NBER conference on industrial organization held at the University of California, Berkeley. Only I can be blamed for defects that remain, of course.

I. Introduction

In a classic essay, Walter Oi [1971] considers a profit-maximizing Disneyland monopoly setting prices for admission to the park and for individual rides. The park's costs are assumed to vary only with the number of rides taken, while households derive utility only from rides, not from being in the park. If all households were identical, the optimal policy would be to sell rides at marginal cost and to set the admission fee so as to convert all the consumers' surplus to profit. If consumers' tastes differ but the park's management knows the relation in the population between rides demanded and individual surplus, Oi indicates that the two prices can be set in a discriminatory fashion to enhance the park's profits. Our concern here is with this sort of imperfect discrimination as a response to heterogeneity of potential buyers.

As Oi argues, situations of this general type are common. Polaroid film and Polaroid cameras must be used together, for instance, and Polaroid had a monopoly on instant photography for a considerable time. If buyer utility depended only on the number of instant photos taken, not on camera ownership, then Polaroid's pricing problem during its monopoly period differed from that analyzed by Oi only in that cameras cost something to produce, while Oi associated zero cost with admissions. Mitchell [1978] uses essentially the Oi demand structure to analyze the demand for local telephone service as a function of connection and per-call charges. Metering schemes of various sorts, either literally involving meters or operating through tying arrangements, may be described in similar terms.[1]

Since the Oi paper appeared, a good deal has been written about two-part tariffs. Most of that writing has focused on the use of such tariffs by welfare-maximizing legal monopolies.[2] This essay, on the other hand, is concerned mainly with the use of two-part pricing arrangments by profit-maximizing firms, though a number of results scattered throughout the welfare-maximizing literature are presented in a unified framework. In both situations, more complex multi-part pricing schemes are generally more attractive in principle than the arrangements studied here.[3] Such schemes do not appear to be especially common outside the public utility sector, however. This may be due to legal (Robinson-Patman) problems associated with quantity discounts or because such discounts increase the difficulty of preventing resale.

In the general case considered here and in all other studies of two-part pricing of which I am aware, the monopolist is the sole supplier of two inputs used in the production of some ultimate commodity, which in turn yields either utility or profit. These inputs can be used in variable proportions; in fact, discrimination turns on differences in the proportions chosen by different sorts of buyers.[4] One unit of the fixed input (admission to the amusement park, a Polaroid camera, access to the telephone network) must be purchased in order to produce any of the ultimate commodity, but as much as is desired can be produced with only one unit of the fixed input. Let

R = the price charged by the monopoly for the fixed input,

F = the (constant) marginal cost of producing the fixed input,

and

N = unit sales of the fixed input.

On the other hand, the output of the ultimate commodity varies directly with the amount of the variable input (merry-go-round rides, Polaroid film, telephone calls placed) used. Let

P = the price charged by the monopoly for the variable input,

V = the (constant) marginal cost of producing the variable input,

and

Q = unit sales of the variable input.

The assumption of constant marginal cost serves mainly to simplify notation; its relaxation adds little new insight. In some situations the fixed input may be merely the monopolist's permission to buy the variable input, while in others it may be the use of a specialized piece of equipment.

Because the monopoly's two products are complements, both Q and N generally depend on both R and P. The technology described in the preceding paragraph restricts the form of that dependence. Though the existing literature has made some use of these restrictions, this essay contains what seems to be the first general description of the structure of the demand functions Q(R,P) and N(R,P).

Three cases are considered here, two of which do not seem to have been previously studied in detail in this context. Section 2 follows Oi and makes the assumption that all customers are households and that income effects are zero. This simplest situation, referred to as _Case I_ below, provides a useful benchmark. Section 3 investigates the implications of two more complex situations. In _Case II_, customers are households, but non-zero income effects are allowed. Only Ng and Weisser [1974] seem to have considered this situation in the present context.[5] Some of the classic tying contracts and metering schemes involve industrial equipment sold or leased to competing firms.[6] Accordingly, Section 3 also deals with Case III, in which all customers are competitive firms selling in the same market and able collectively to alter the price in that market. (If all customers are competitive firms facing fixed output prices, Section 3 indicates that Case I applies.) Case III seems not to have been analyzed before in this context; our treatment is based on the work of Panzar and Willig [1978].

Cases II and III differ from Case I in that changes in R affect the demand for the variable input by infra-marginal buyers, those whose decision to purchase the fixed input is unaffected. In Case II this occurs because of straightforward income effects, while in Case III infra-marginal firms respond to changes in the downstream output price induced by entry and exit of marginal firms.

Sections 4 and 5 consider the equilibrium conditions for a monopoly controlling both inputs. This situation is naturally and instructively imbedded in the standard problem of maximizing aggregate Marshallian welfare subject to a profit constraint. The optimality conditions are expressed in terms that permit intuitive interpretations. Because Case I is noticeably simpler than the other two, Section 4 concentrates on it. Section 5 then considers the implications of income effects in Case II and output price effects in Case III. It turns out that normal consumer goods (II) resemble inferior inputs (III) here, while normal inputs resemble inferior goods. The results of Oi [1971] and others on the relation between P and V are generalized, as is the Ng and Weisser [1974] analysis of conditions under which R must be non-negative. It is shown that the Leland-Meyer [1976] proof that if $F = 0$ in Case I, R must be positive does not generalize to Cases II and III. Moreover, it is shown that if F is positive (as in the Polaroid and telephone examples), it may well be optimal to set R less than F, to sell the fixed input below cost and earn all profit on the variable input.[7] Since P = V and R > F is optimal when buyers are identical households, it is perhaps not surprising that buyer diversity is one factor tending to make R < F optimal.

Section 6 compares the various Case I equilibria described in Section 4 under the assumption of "slight" monopoly power, which ensures that the equilibria are "close". Section 7 extends this analysis to Cases II and III. As one might expect by analogy with classic second- and third-degree price discrimination, the output and welfare implications of a movement from monopoly over one input alone to a two-part pricing monopoly depend on the demand structure.[8] Under plausible conditions, it is shown that if a monopoly over the fixed (variable) input is extended to both inputs by a tying contract or equivalent pricing arrangement, aggregate welfare and total sales of the fixed (variable) input both rise. It is also argued that equity considerations do not provide much justification for antitrust hostility toward tying contracts. On the other hand, two-part pricing that involves persistent sales below cost of one of the inputs is shown generally to lead to drops in Marshallian welfare.

II. Demand Relationships: Case I

We assume a continuous distribution of household types, indexed by θ. Without loss of generality, suppose that $0 \leq \theta \leq 1$. One might associate θ with household income in some situations, but there is no need to make that association explicit in general. Let

$S(P,\Theta)$ = Marshallian surplus enjoyed by a household of type Θ when $R = 0$.

$q(P,\Theta)$ = the quantity of the variable input purchased by a household of type Θ when $R = 0$.

The quantity S is the area between the ordinary individual demand curve and the price line. It follows from the geometry of the situation that $\partial S(P,\Theta)/\partial P = -q(P,\Theta)$.

Assume that S is increasing in Θ in the relevant region. That is, assume that over the range of values of P considered, the ordering of households by surplus enjoyed does not change. This certainly holds if demand curves corresponding to different Θ's never cross, as Littlechild [1975], Leland and Meyer [1976], and others have assumed, but it is somewhat more general. If the individual demand curves don't cross or, more generally, if those who buy more of the variable input enjoy the larger surpluses, it must be that $\partial q(P,\Theta)/\partial\Theta > 0$. This is referred to as the <u>direct case</u>, with $\partial q(P,\Theta)/\partial\Theta$ negative in the <u>inverse case</u>. It is not necessary that either of these occur, of course, since the sign of $\partial q/\partial\Theta$ may vary with Θ for any or all values of P.

A household of type Θ will rationally purchase one unit of the fixed input if and only if $S(P,\Theta) \geq R$. If and only if the fixed input is purchased, the household will buy $q(P,\Theta)$ units of the variable input. The main role of the assumption that S is increasing in Θ is to ensure that the marginal customer type, defined implicitly by

$$S(P,\hat{\Theta}) = R \tag{1}$$

is unique. If this assumption is relaxed so that (1) may have multiple solutions, much of what follows goes through with weighted averages of marginal customers' purchases appearing in the obvious places; see Ng and Weisser [1974, Appendix] for a discussion.

The analysis below deals mainly with situations in which some but not all households buy the fixed input, so that $0 < \hat{\theta} < 1$. This seems the most generally relevant case for applications, and once it is understood the analysis of either excluded polar case is relatively straightforward. As long as S is smooth, differentiation of (1) totally yields

$$\hat{\theta}_R = 1/S_\theta(P,\hat{\theta}) > 0, \text{ and} \tag{2a}$$

$$\hat{\theta}_P = q(P,\hat{\theta})/S_\theta(P,\hat{\theta}) = \hat{q}\hat{\theta}_R > 0 \tag{2b}$$

where here and in all that follows subscripts indicate partial differentiation and $\hat{q} = q(P,\hat{\theta})$. The relation between these two derivatives arises because the marginal consumer is excluded by lowering his post-purchase real income, and the income effects of a small increase in R of ΔR and a small increase of ΔP in P are equivalent if and only if $\Delta R = \hat{q}\Delta P$. Relations in (2) establish that N is decreasing in both R and P in Case I.

If $m(\theta)$ is the density function of household types, N and Q are given by the following integrals:

$$N(R,P) = \int_{\hat{\theta}(R,P)}^{1} m(\theta)d\theta, \qquad Q(R,P) = \int_{\hat{\theta}(R,P)}^{1} q(P,\theta)m(\theta)d\theta \tag{3}$$

With income effects assumed away in Case I, the following quantity measures only the infra-marginal substitution effect:

$$\sigma = \int_{\hat{\theta}(R,P)}^{1} q_P(P,\theta)m(\theta)d\theta \tag{4}$$

Differentiation of (3), substitution from (4), and use of (2) yield

Proposition 1: In Case I, the demand derivatives are related as follows: (a) $N_P = \hat{q}N_R = Q_R$, and (b) $Q_P = \hat{q}N_P + \sigma$.[9]

The first quality in (a) follows directly from (2); it simply re-states the fact that a small variable input price increase of ΔP has the same exclusion effect as an increase of $\hat{q}\Delta P$ in the fixed input price. That is, they impose the same post-purchase real income loss on the marginal buyer. The second equality in (a) follows because increases in R lower sales of the variable input only by excluding marginal buyers of the fixed input; with no income effects, R is otherwise irrelevant. Part (b) shows that the impact of P on Q is made up of an exclusion effect of the type just discussed and an infra-marginal substitution effect. Note that $\hat{q}$, σ, and one of the four demand derivatives serve to determine the other three derivatives. The (standard) assumptions made here thus impose considerable structure on demand.

III. Demand Relationships: Cases II and III

In order to add income effects to the model just considered so as to move to Case II, it is convenient to define

$I(P,Y,\theta)$ = the indirect utility function of a household of type θ with income Y able to purchase the variable input at price P.

A gross surplus measure can be defined as follows, where $Y(\theta)$, is the pre-purchase income of a household of type θ.

$$S(R,P,\theta) = I[P,Y(\theta)-R,\theta] - I[\infty,Y(\theta),\theta] + R$$

The first term on the right gives the utility enjoyed if R is spent on the fixed input and the variable input is then purchased at P, while the second term gives the utility level corresponding to exclusion from this market. The difference between them gives the surplus, in units of utility, enjoyed by a household purchasing the variable input.

Let us assume that there exists a set of cardinal indirect utility functions, consistent with household preferences, such that $S_\theta(R,P,\theta)$ is positive in the relevant region, as before. Then it is clear that $S(R,P,\hat{\theta}) = R$ uniquely defines the marginal household type. This condition itself is not dependent on choice of a cardinal representation. Let us also assume in all that follows that I has been normalized so that $I_Y[P,Y(\theta)-R,\theta] = 1$ for all $\theta > \hat{\theta}$. This amounts to assuming that income is distributed optimally in normative analysis; here it serves mainly to simplify notation.[10] With this assumption, differentiation and Roy's Theorem yield[11]

$$\hat{\theta}_R = 1/S_\theta(R,P,\hat{\theta}) > 0 \tag{5a}$$

$$\hat{\theta}_P = -\, I_P[P,Y(\hat{\theta})-R,\hat{\theta}]/S_\theta(R,P,\hat{\theta}) = \hat{q}\hat{\theta}_R > 0 \tag{5b}$$

The only substantive difference between equations (2) and (5) is that in (5b), $\hat{q}$ is a function of R as well as of P and θ. It is also easy to verify that $S_R(R,P,\theta) = 0$ and $S_P(R,P,\theta) = -q(R,P,\theta)$. Thus for purposes of demand analysis Cases I and II differ mainly in that infra-marginal households' demands vary with R in Case II.

Two different infra-marginal income effects appear in this Case:

$$\eta = \int_{\hat{\theta}(R,P)}^{1} q_Y(\theta)m(\theta)d\theta, \qquad \eta' = \int_{\hat{\theta}(R,P)}^{1} q(\theta)q_Y(\theta)m(\theta)d\theta \tag{6}$$

where $q_Y(\theta)$ is the derivative of the demand of customer type θ with respect to income. If both η and η' are positive (negative) in the relevant region, the ultimate commodity here is called a <u>normal good</u> (<u>inferior good</u>) in what follows. (In general, of course, η and η' may be of different signs and may change sign as R and P change.) With the own-price derivatives in (4) now explicitly interpreted as utility-constant compensated derivatives, straightforward differentiation and use of the Slutsky decomposition yield the required generalization of Proposition 1:

<u>Proposition 2</u>: In Case II, the demand derivatives are related as follows: (a) $N_P = \hat{q}N_R$, (b) $Q_R = \hat{q}N_R - \eta$, and (c) $Q_P = \hat{q}N_P + \sigma - \eta'$.

The intuition behind part (a) is exactly as before: the exclusion effect depends on the amount by which the post-purchase real income of the marginal buyer is lowered. Part (b) shows that the impact of R on Q consists of an exclusion effect, as in Proposition 1, plus an infra-marginal <u>income effect</u>. For a normal good, the real income fall caused by a rise in R serves to reduce infra-marginal demand for the variable input. Similarly, part (c) splits the impact of P on Q into an exclusion effect, an infra-marginal substitution effect, and another infra-marginal income effect.

Let us now turn to Case III, in which all customers are competing firms. Assume that all produce a single ultimate output, with $x(\theta)$ the supply of a firm of type θ. Let Z be the price of the ultimate output, and let D(Z) be the corresponding market demand function. (In this context, the activities of firms producing this same output but not using the fixed input considered here are assumed to be reflected in D(Z), which is thus a net demand curve facing the relevant set of firms.) It is natural to define

$S(P,Z,\theta)$ = the (economic) profit earned by a firm of type θ with input price P, output price Z, and $R = 0$.

The fixed input will be purchased by any firm if and only if $S \geq R$, exactly as above. Let us continue to assume that it is possible to order customers so that S is increasing in θ everywhere relevant. By Hotelling's Lemma,[12]

$$S_P(P,Z,\theta) = -q(P,Z,\theta), \tag{7a}$$

$$S_Z(P,Z,\theta) = x(P,Z,\theta), \text{ and} \tag{7b}$$

$$S_{PZ}(P,Z,\theta) = -q_Z(P,Z,\theta) = x_P(P,Z,\theta) \tag{7c}$$

If Z is fixed, it should be clear that the analysis of Case I applies exactly. Then (7a) is the necessary derivative property, $S \geq R$ is the test for purchase, and infra-marginal demand for the variable input is independent of R. (If a firm decides to operate at all, R becomes a sunk cost.)

If the relevant set of firms can affect the output price, the situation becomes more complicated. One must consider Z as a function of R and P and then exploit relations (7). The highlights of the development are given in the Appendix. Two infra-marginal output price effects arise here:

$$\zeta^R = (Z_R)\int_{\hat{\theta}(R,P)}^{1} q_Z(P,Z,\theta)m(\theta)d\theta, \quad \zeta^P = (Z_P/Z_R)\zeta^R \tag{8}$$

If we are dealing with a <u>normal input</u> (<u>inferior input</u>), q_Z is positive (negative).[13] It is shown in the Appendix that increases in R always raise Z, so that ζ^R is positive (negative) for a normal (inferior) input. In the normal case, an increase in P serves both to exclude marginal suppliers and to reduce the supply of infra-marginal firms. Both effects increase Z, so that ζ^P is also positive for a normal input. As Panzar and Willig [1978] stress, however, increases in P <u>raise</u> the optimal outputs of infra-marginal firms in the inferior case, and Z_P and ζ^P have indeterminate signs in that case.

Differentiation of (3), treating q as a function of P, Z(R,P), and θ, and exploitation of the relations developed in the Appendix yield the demand relationships for Case III:

<u>Proposition 3</u>: In Case III, the demand derivatives are related as follows: (a) $N_P = Q_R = \hat{q}N_R + \zeta^R$, and (b) $Q_P = \hat{q}N_P + \sigma + \zeta^P$.

In part (b), σ is defined by (4) with q_p understood to be also a function of Z. Thus σ is a negative infra-marginal substitution term as in Cases I and II, except that output price is held constant here, not utility or profit.[14]

From Proposition 1, $N_P = Q_R$ in the absence of any effects on infra-marginal demand. This equality is not generally valid in Case II because R may affect directly the infra-marginal demand for the variable input, while P clearly can never affect the infra-marginal demand for the fixed input. This equality <u>is</u> valid in Case III, as Proposition 3 indicates, where R does not have any such direct infra-marginal effect. Since $N_P = Q_R$ with Z constant, since there are no direct infra-marginal effects involved, and since only induced changes in Z need to be considered in both cases, it is perhaps not too surprising that $N_P = Q_R$ when Z is allowed to change to restore full equilibrium.

The second equality in part (a) of Proposition 3 decomposes the impact of R on Q into an exclusion effect and an output price effect. For normal inputs, the latter tends to offset the former. As suppliers drop out of the market, the output price tends to rise, increasing the input demands of infra-marginal firms. Note that in Case II, the income effect adds to the exclusion effect for normal goods. In the normal case, part (a) also indicates that a small price increase of ΔP has less impact on N than a small increase in R of $\hat{q}(\Delta P)$. This is because the increase in P lowers infra-marginal supplies for all values of Z, thus tending to raise Z and offset the exclusion effect, while changes in R have no direct infra-marginal impact.

A comparison of parts (b) of Propositions 2 and 3 shows that a normal input also behaves like an inferior consumer good when P changes. In both cases, increases in P that tend to exclude marginal buyers are offset to some extent by increases in infra-marginal demand. Since ζ^P cannot be signed in general for inferior inputs, the analogy between them and normal consumer goods is less exact.

IV. Equilibria and Optima: Case I

The standard distribution-free welfare measure used in all that follows:[15]

$$W(R,P) = \int_{Z(R,P)}^{1} D(x)\,dx + \int_{\hat{\theta}(R,P)}^{1} \{S(P,..,\theta) - F + (P-V)q(P,..,\theta)\} m(\theta)\,d\theta, \quad (9)$$

where Z is a constant in Cases I and II, and the unlisted arguments in the second integrand are (I) none, (II) R, and (III) Z(R,P). Note that in Case III, the Marshallian measure is applied to the ultimate commodity.

It is straightforward to verify that in all three Cases, the first-order conditions for maximization of W can be written

$$W_R = (R-F)N_R(R,P) + (P-V)Q_R(R,P) = 0 \tag{10a}$$

$$W_P = (R-F)N_P(R,P) + (P-V)Q_P(R,P) = 0 \tag{10b}$$

These are obviously satisfied by marginal cost pricing, by setting R = F and P = V.[16] If one of the two prices is fixed above cost and profitability is not a binding constraint, conditions (10) show that the other price should be set below cost. The two goods are complements, so that second-best pricing involves offsetting exogenous distortions, not the sort of distortion matching one thinks of for substitutes. The second-order conditions that ensure that marginal cost pricing maximizes W are the following:

$$N_R < 0,\ Q_P < 0, \text{ and } N_R Q_P - (N_P)^2 > 0 \tag{11}$$

By Proposition 1, the third of these is satisfied in Case I because σ is negative, and the first two also hold for all non-negative R and P.

Figure 1 shows the relevant geometry near the marginal-cost-pricing point, w. The locus $R^W(P)$ is defined by (10a), $P^W(R)$ by (10b). Both are negatively sloped at w, and the second-order conditions ensure that the former intersects the latter from below. The ellipse is an iso-W curve.

It is convenient to express some of the conditions developed below in terms of the following demand elasticities:

$$e^P = - O_P(P/Q), \qquad e^R = - Q_R(R/Q), \tag{12a}$$

$$\mu^P = - N_P(P/N), \qquad \mu^R = - N_R(R/N), \tag{12b}$$

$$\bar{e}^P = - \bar{Q}_P(P/Q) = - [Q_P - (N_P/N_R)Q_R](P/Q), \tag{12c}$$

$$\bar{\mu}^R = - \bar{N}_R(R/N) = - [N_R - (Q_R/Q_P)N_P](R/N). \tag{12d}$$

Note that $\bar{Q}_P$ is the compensated derivative of Q with respect to P, with R changed so as to hold N constant. Similarly, $\bar{N}_R$ is the Q-constant derivative of N with respect to R. In general, $\bar{N}_R = \bar{Q}_R(N_R/Q_P)$. In Case I, $\bar{Q}_P = \sigma$, so that $\bar{Q}_P$ and $\bar{N}_R$ are negative, and both $\bar{e}^P$ and $\bar{\mu}^R$ are positive.

If a firm monopolizes the fixed input only and must treat the price of the variable input as exogenous, it will choose R so as to maximize

$$\Pi^f(R,P) = (R - F)N(R,P).$$

Let the solution be $R^f(P)$. If the variable input is produced competitively, the fixed-input monopoly will operate at a point like f in Figure 1. The slope of the $R^f(P)$ curve cannot in general be signed at f or elsewhere.

Similarly, a variable-input monopolist that must take R as given will choose P to maximize

$$\Pi^v(R,P) = (P - V)Q(R,P).$$

If R = F, operation will be at a point like v in Figure 1. As above, the slope of the firm's optimal response function, $P^v(R)$, cannot be signed.

A firm with control over both inputs, whether acquired through tying arrangements or by other means, will seek to maximize total profit, $\Pi \equiv \Pi^f + \Pi^v$. It is easy to imbed this in the problem of maximizing W subject to a lower bound constraint on Π, the problem with which most of the literature on two-part tariffs has been concerned. Solutions to the latter problem correspond to stationary points of

$$\Psi(R,P) = \omega\Pi(R,P) + (1-\omega)W(R,P), \tag{13}$$

where $0 \leq \omega \leq 1$.[17] Values of ω near or equal to zero are generated by weak or non-binding profit constraints, while values near unity are associated with tight constraints. As ω moves from zero to unity, the solution prices move along a locus of tangencies between iso-W and iso-Π curves from a point like w in Figure 1 to a profit-maximizing point like π.

In what follows, solutions corresponding to strictly positive values of ω are called <u>two-part optima</u>, with the <u>two-part profit maximum</u> point being an extreme point in this set. The excluded marginal-cost-pricing point w is called the <u>unconstrained welfare maximum</u>.

Differentiating (13) and using equations (10), one obtains the two basic first-order contitions:

$$\Psi_R = \omega N + [R-F+(P-V)\hat{q}]N_R + (P-V)[Q_R - \hat{q}N_R] = 0, \tag{14a}$$

$$\Psi_P = \omega Q + [R-F+(P-V)\hat{q}]N_P + (P-V)[Q_P - \hat{q}N_P] = 0. \tag{14b}$$

The first term in both these conditions is ω times the direct revenue gain on infra-marginal sales caused by an increase in R or P. This is a pure transfer from buyers to the seller; it is valued here exactly to the extent that profits matter. In both of equations (14), the second term is the net profit on sales to the marginal customer, times the derivative of N with respect to the decision variable. The marginal customer retains no surplus, since $S(\hat{\theta}) = R$, so that the monopolist's profit gain on sales to that customer exactly equals the net social welfare gain. The third term in both equations is (P-V) times the change in the infra-marginal demand for the variable input induced by an increase in R or P. (Note that the exclusion effect is subtracted from the net effect.) This is again the relevant quantity for <u>both</u> welfare and profit computations.

Suppose $\omega = 1$ in conditions (14), so that they describe the two-part profit maximum. Let $R^{\pi}(P)$ be the result of solving (14a) for R as a function of P, and let $P^{\pi}(R)$ be similarly obtained from (14b). Figure 1 illustrates the relation between $R^{\pi}(P)$ and the fixed-input monopoly locus, $R^f(P)$. It is easy to show that R^{π} intersects R^f from above at point f and that they intersect only there. Thus if P exceeds V, R must be held below the level that would maximize profit on the fixed input alone, since increases in R tend to reduce Q and thus to reduce profits on the variable input. Exactly symmetrically, $P^{\pi}(R)$ intersects $P^V(R)$ from the left at the variable-input monopoly point, v. Finally, the second-order conditions for profit maximization imply that the R^{π} and P^{π} loci intersect as shown at the two-part profit maximum, point π.

In Case I, Proposition 1 establishes that the right-most bracketed term in (14a) is zero, while the corresponding term in (14b) is equal to the negative substitution effect, σ. From the first of these relations we have

Proposition 4: In Case I, the net profit from sales to the marginal customer, $[R-F+(P-V)\hat{q}]$, is strictly positive at a two-part optimum.

Figure 1 is drawn so that profit maximization involves selling both fixed and variable inputs at prices that exceed the corresponding marginal costs. This is not always optimal, however, as we now show. Eliminating the bracketed term for profit on the marginal customer from equations (14) and solving, one obtains

$$(P-V)/P = [\omega/\bar{e}^P][1 - (N_P/\bar{q}N_R)], \tag{15}$$

where $\bar{q} = Q(R,P)/N(R,P)$ is the amount of the variable input used by the average buyer. For profit maximization, $\omega = 1$. The condition for maximization of Π^V, profit on the variable input alone, replaces the right-hand side of (15) with $(1/e^P)$. Since $\bar{e}^P$ is positive in Case I, Proposition 1 establishes [17]

Proposition 5: In Case I, at a two-part optimum, $(P-V)$ has the sign $(\bar{q}-\hat{q})$.

In the direct case, if $P > V$, the largest buyers, who have the largest gross surpluses, make the largest contribution to profit. In the inverse case, setting $P < V$ creates the same situation, since then those with the largest q's have the smallest surpluses and contribute least to profit. Though the general message of Proposition 5 is thus no surprise, it is interesting to learn that $(\bar{q}-\hat{q})$ is a sufficient statistic for general demand patterns.

The relation between R and F is less simply described, as intuition might suggest. After all, only differences in buyers' q's can be observed or exploited by the seller, and only the variable input price translates these demand differences into differential contributions to profit. The price of the fixed input cannot be used to discriminate; it plays a sort of residual role here.

Ng and Weisser [1974] assumed F = 0 in a situation like our Case II and argued that the optimal R must be non-negative.[19] Here we can easily prove

Proposition 6 (Ng-Weisser): In Case I, $R \geq 0$ at a two-part optimum.

If the optimal R were negative, customers could receive a subsidy even if they bought none of the variable input. All would elect to receive the subsidy, and small changes in R or P would then have no effect on N. Since $Q_R = \hat{q}N_R$ in Case I, if $N_R = 0$ condition (14a) would reduce to $\omega N = 0$. This clearly cannot hold at a two-part optimum, so the optimal R cannot be negative.

Leland and Meyer [1976] showed that if F = 0 in Case I, R must be strictly positive. Their argument generalizes easily to non-negative F:

Proposition 7 (Leland-Meyer): In Case I, if F is non-negative, $R > 0$ at a two-part optimum.

If the optimal R is zero, condition (14a) may be written

$$\Psi_R = \omega N - FN_R + (P - V)Q_R = 0.$$

Since N_R is negative, Proposition 7 is established by showing that $Q_R = 0$, thus showing that R = 0 is incompatible with the necessary condition (14a). The argument has two steps. (a) If R = 0 in Case I, $\hat{q}$ must equal zero also. With downward sloping demand curves, positive purchases must imply positive surplus, but with R = 0, $S(\hat{\theta}) = 0$ also, and this means $\hat{q} = 0$. (b) From Proposition 1, $\hat{q} = 0$ implies $Q_R = 0$, since R affects Q only by exclusion.

In order to investigate the general relation between R and F when both are non-negative, eliminate (P-V) from equations (14) and re-arrange to obtain

$$(R-F)/R = [\omega/\bar{\mu}^R][(\bar{q}Q_R/Q_P) - 1]. \tag{16}$$

Maximization of Π^f, profit from sales of the fixed input alone, requires setting (R-F)/R equal to $(1/\mu^R)$. As with condition (15), the ordinary elasticity is replaced here by the compensated elasticity. Note also that conditions (15) and (16) are perfectly general; they do not depend on any of the structure presented in Sections 2 and 3. Since $\bar{\mu}^R$ is positive, Proposition 1 can be used directly with condition (16) to establish

<u>Proposition 8</u>: In Case I, at a two-part optimum (R-F) has the sign of $[(\bar{q} - \hat{q})Q_R - \sigma]$.

If all buyers are identical, $\bar{q} = \hat{q}$, and R > F, as Oi [1971] observed. If $\bar{q} < \hat{q}$, Proposition 5 indicates that P is optimally below

V, so it is no surprise that R must exceed F. On the other hand, if the marginal customer buys more than average, as in the direct case, the sign of (R-F) is in general ambiguous. In the direct case, Propositions 5 and 8 indicate that greater buyer diversity, as measured by a larger value of $(\hat{q} - \bar{q})$ tends to lower R and to raise P. The stronger is the infra-marginal substitution effect, as measured by the absolute value of ϑ, the greater demand response to increases in the price of the variable input, and the more likely it is to be optimal to earn profit on the fixed input by setting R above F.

In Case I, if $\bar{q} < \hat{q}$, the variable input is always sold below cost at a two-part optimum, and profits are earned on sales of the fixed input. If buyers are identical, the variable input is sold at cost, and R is optimally set above F. This difference derives from the infra-marginal substitution effect associated only with the variable input. If $\bar{q} > \hat{q}$, the variable input is sold above cost, and the sign of (R-F) is ambiguous. All else equal, the greater the gap between $\bar{q}$ and $\hat{q}$, the larger will be (P-V) and the smaller will be (R-F). If one thinks that the direct case, with $\bar{q} > \hat{q}$, is typical and that considerable buyer diversity is also the norm, this analysis suggests that pricing strategies that involve "giving away the razor amd making money on the blades" are more commonly optimal than policies in which most profits are earned on sales of the fixed input. Casual empiricism supports this suggestion.[20]

V. Equilibria and Optima: Cases II and III

We now employ Propositions 2 and 3 to consider extensions of Propositions 4 - 8 to situations with income effects (Case II) or output price effects (Case III). Since equations (15) and (16) hold in general, it is important at the outset to establish the sign of the compensated derivative $\bar{Q}_P$ and thereby that of the compensated elasticity $\bar{e}^P$, along with the signs of $\bar{N}_R$ and $\bar{\mu}^R$. First, we assume that conditions (11), which ensure that marginal cost pricing maximizes welfare, hold globally. In Case III, this suffices to prove that $\bar{Q}_P$ is everywhere negative. In Case II, these second-order conditions rule out Giffen goods by requiring $(\sigma - \eta')$ to be negative. Unfortunately, this implies only that $\bar{Q}_P < \hat{q}\eta$. We thus have $\bar{Q}_P$ negative for inferior goods but of ambiguous sign for normal goods. To rule out a variety of pathologies induced by strong income effects, it appears necessary to assume explicitly that $\bar{Q}_P$ is everywhere negative in Case II also.[21] It then follows that $\bar{N}_R < 0$ and $\bar{\mu}^R > 0$ for all non-negative R and P in both cases.

In Case II, the right-most bracketed term in (14a) is equal to $(-\eta)$, while in Case III it equals (ζ^R). Proposition 4 thus generalizes to

Proposition 4': The net profit from sales to the marginal customer, $[R-F+(P-V)\hat{q}]$, is strictly positive at a two part optimum in Case II if $(P-V)\eta < 0$ and in Case III if $(P-V)\zeta^R > 0$.

If $P > V$, the sign of marginal profit is in general ambiguous for normal goods in Case II and for inferior inputs in Case III. In both these situations, reductions in R stimulate infra-marginal demand, and if $P > V$ it may be optimal even for a profit-maximizer to lower R to the point where money is lost on the marginal customer.

Application of Propositions 2 and 3 to the variable input markup condition, equation (15), establishes

Proposition 5': At a two-part optimum, (P-V) has the sign of $(\bar{q} - \hat{q})$ in Case II and the sign of $[(\bar{q} - \hat{q}) + (-1/N_R)\zeta^R]$ in Case III.

Cases I and II are identical here. In Case III, for normal inputs, it is sufficient but not necessary (as in Cases I and II) that $\bar{q}$ exceed $\hat{q}$ for P to exceed V. To see why, assume for the moment that all purchasing firms are identical and that a profit-maximizer has set $P = V$ and $R > F$, as would be optimal in Cases I and II. As (a) of Proposition 3 indicates, the seller can then lower R by a small amount ΔR, increase P by more than $(\Delta R)/\hat{q}$, and still leave N unchanged. (As was noted below Proposition 3, this occurs because the increase in P lowers infra-marginal supplies of the final output, thus tending to raise Z and N, while the drop in R has no offsetting infra-marginal impact.) As this means that net revenue from all customers is higher, profit is increased. Similarly, in the special case of identical buyers, the monopolist optimally sets P below V if the input is inferior, since this tends to increase Z and thereby raise N.

If income effects are positive, strong enough, and appropriately matched to household differences, it is apparently possible for the optimal R to be negative in Case II.[22] Proposition 6 generalizes as follows:

<u>Proposition 6'</u>: At a two-part optimum, R is non-negative in Case III. R is also non-negative in Case II if <u>any</u> of the following are satisfied:

$$\text{(a)} \quad \eta \leq 0, \qquad \text{(b)} \quad (\bar{q}-\hat{q})\eta \leq 0, \qquad \text{or (c)} \quad \eta-\eta'/\bar{q} \leq 0.$$

As in the proof of Proposition 6, if R is negative, $N_R = N_P = 0$. In Case III, Proposition 3 establishes that Q_R must then be zero also, and necessary condition (14a) cannot be satisfied. In Case II, Proposition 2 shows that η must be positive in order to satisfy this condition, so that part (a) of Proposition 6' is established. Elimination of $(P-V)/\omega$ between conditions (14) yields the necessary condition

$$Q_P - \bar{q}Q_R = \sigma - \eta' + \bar{q}\eta = \bar{Q}_P + (\bar{q} - \hat{q})\eta = 0.$$

Since σ and $\bar{Q}_P$ are negative, parts (b) and (c) of Proposition 6' follow directly.

It is instructive to interpret the negatives of (a) - (c) in Proposition 6' as necessary conditions for the optimal R to be negative in Case II. First, part (a) indicates that the good must be normal, so that a reduction in R stimulates infra-marginal demand for the variable input. Second, $\bar{q}$ must exceed $\hat{q}$, so that (by Proposition 5') it is optimal to set P above V and earn a profit on that demand. Finally, the q-weighted average income effect, which arises when P is changed, must be less than the unweighted average that gives the infra-marginal response to changes in R.

Perhaps not surprisingly, the Leland-Meyer [1976] argument used to establish that R is positive in Case I does not go through in either Case II or Case III. That argument, presented to establish Proposition 7, above, has two main steps: (a) $R=0 \Rightarrow \hat{q}=0$, and (b) $\hat{q}=0 \Rightarrow Q_R=0$. Step (a) fails in Case III, since a firm can have S = R = 0 and still be using strictly positive amounts of all inputs. Step (b) fails in Case II because of income effects on infra-marginal demand. Thus even if F is positive in these Cases, it may be optimal to set R = 0, to use ordinary uniform pricing even when a two-part tariff is feasible.[23]

Finally, use of condition (16) and Propositions 2 and 3 establishes the following generalization of Proposition 8:

<u>Proposition 8'</u>: At a two-part optimum, (R-F) has the sign of $[(\bar{q}-\hat{q})Q_R - \bar{Q}_P]$

in Case II and the sign of $[(\bar{q}-\hat{q})Q_R - \bar{Q}_P - (Q_R/N_R)\zeta^R]$ in Case III.

The qualitative implications for Case II are exactly the same as those discussed below Proposition 8 for Case I. (Recall that $\bar{Q}_P = \sigma$ in Case I.) Just as a large value of ζ^R was shown in Proposition 5' to make it more likely that (P-V) is positive, so it is shown here to make it more likely that (R-F) is negative. The intuition is the same in both situations.

VI. Local Comparisons: Case I

This Section and the next are devoted to pairwise comparisons among the four equilibria depicted in Figure 1: w, f, v, and π.[24] Two of these comparisons are straightforward: the move from w to f or from w to v is an ordinary move from competition to monopoly under our constant cost assumption. Both transitions involve drops in N, Q, and W. The comparison between f and v involves two partial monopoly positions; it is neither especially tractable nor especially interesting.

The three comparisons that remain are more interesting: w versus π, f versus π, and v versus π. Since π is not quite an ordinary monopoly equilibrium point, its relation to w is not obvious in all respects. The final two comparisons are directly relevant to public policy toward tying contracts and sales below cost. If a firm has a monopoly only over some specialized machine and acts as an ordinary monopolist, it will operate at a point like f in Figure 1. If it can move to π either by selling the variable input below cost so as to raise the optimal rental or (as drawn) by requiring buyers of the fixed input to purchase the variable input from it at a price above V, it can raise its profits. Both sorts of actions can give the firm serious antitrust problems; our aim here is to investigate the economic basis of judicial hostility. Similarly, a firm operating as a variable input monopolist at a point like v in Figure 1 will seek to move to π either by selling the fixed input below cost so as to stimulate demand for the variable input or (as drawn) by requiring purchase of the fixed input at a price above cost. The issue is whether on economic grounds such actions should be applauded or attacked when the model developed here is relevant.[25]

A general equity point arises in all three comparisons. By assumption, consumer's surplus or producer's profits are increasing in θ in all three Cases. Similarly, it follows directly from the analysis in Sections 2 and 3 that $\partial(S_\theta)/\partial P = -\partial q/\partial\theta$. In the direct case, q rises with θ, so that increases in P tend to equalize surplus or profits across buyers, as Figure 2 illustrates. In Case III, this means that small firms are hurt less than large firms, relatively, by policies involving larger values of P. While this might generally be viewed as progressive incidence, the corresponding equalization of consumers' surpluses in Cases I and II may be progressive or regressive, depending on the real incomes of the households involved.[26] Similarly, in the inverse case increases in P increase the inequality of S across buyers; see Figure 2 for an illustration. In what follows, these sorts of changes are termed equalizing and inequalizing, respectively.

Except for w, all equilibria shown in Figure 1 are monopoly points, and their characteristics are dependent upon the values of a number of demand elasticities. The usual problem encountered in comparing monopoly equilibria, the inability to say much *a priori* about changes in those elasticities between distinct points, is present here. One might deal with this problem by working with arbitrary but tractable functional forms for $m(\theta)$ and $S(P,Z,\theta)$, then hoping that the results obtained are valid for a large set of other functions as well. The approach taken here is instead to confine the analysis to comparisons of equilibria that are sufficiently close that first-order approximations can be employed in comparative analysis. This yields results that do not depend on complete descriptions of functional forms, but one can only hope that these have some value when equilibria are a non-negligible distance apart.

Let us begin with the comparison of w, the unconstrained welfare maximum, and some two-part equilibrium "nearby" on the rest of the $w\pi$ locus. The comparison of w and π is a special case of this. It is obvious that a move away from w must lower welfare. The first-order approximations to changes in N and Q are simply

$$\Delta Q = Q_R(\Delta R) + Q_P(\Delta P), \text{ and } \Delta N = N_R(\Delta R) + N_P(\Delta P). \tag{17}$$

We assume that the two points being compared are sufficiently close that (17) can be used along with the tangent to the $w\pi$ locus to evaluate changes in Q and N. Differentiating conditions (14) totally with respect to ω and evaluating the resultant system at $\omega=0$, one obtains

$$dR/d\omega = K(QN_P - NQ_P), \text{ and } dP/d\omega = K(NN_P - QN_R). \tag{18}$$

where K is a positive constant by the second-order conditions (11). Note that $dP/d\omega$ has the sign of $(\bar{q} - \hat{q})$ in Case I, as it should.

Assuming that the changes in (17) are proportional to the derivatives in (18), substitution yields

$$\Delta Q \overset{s}{=} Q(Q_R N_P - Q_P N_R) + NQ_P(N_P - Q_R), \tag{19a}$$

$$\Delta N \overset{s}{=} N[(N_P)^2 - N_R Q_P], \tag{19b}$$

where "$\overset{s}{=}$" is shorthand for "has the same sign as," and all derivatives are evaluated at the unconstrained welfare maximum. Application of Proposition 1 and conditions (11) establishes

Proposition 9: In Case I, if the unconstrained welfare maximum and a two-part optimum are sufficiently close, a move from the former to the latter lowers Q and N. Such a move is equalizing in both the direct and inverse cases.

In comparisons not involving the unconstrained welfare maximum, the direction of movement of W is a priori uncertain. From the first equalities in conditions (10) and equations (17), we obtain the relevant first-order approximation:

$$\Delta W = (R - F)\Delta N + (P - V)\Delta Q. \tag{20}$$

Note that at the fixed-input (variable-input) monopoly point, the second (first) term is zero, and ΔW has the same sign as ΔQ (ΔN).

Now consider movements from the fixed input monopoly point, point f in Figure 1, to the two-part profit maximum, point π. From Figure 1, movement from f to π can always be accomplished by movement along the $R^{\pi}(P)$ locus, defined by first-order condition (14a) with $\omega = 1$. If the two points are close together, one need only know the slope of that locus at f. Re-writing (14a) in terms of the elasticity μ^R and differentiating, one obtains that slope as

$$(dR/dP)^f = [\mu^R RQ_R - NR\mu^R_P]/[NR\mu^R_R + N\mu^R(\mu^R - 1)],$$

where μ^R_P and μ^R_R are the partial derivatives of μ^R with respect to P and R, respectively, at point f. The second-order conditions for a fixed-input monolopy ensure that the denominator is positive.

Substitution of $(dR/dP)^f(\Delta P)$ for (ΔR) in (17) and substitution into (20) yield

$$\Delta W \stackrel{s}{=} \Delta N \stackrel{s}{=} \{N[N_P\mu_R^R - N_R\mu_P^R] + N_R N_P + N_R\mu^R[Q_R - N_P]\}\Delta P. \tag{21}$$

The corresponding expression for the sign of ΔQ cannot easily be simplified. But use of Proposition 1 allows one to re-write (21) as

$$\Delta W \stackrel{s}{=} \Delta N \stackrel{s}{=} \{\mu_P^R + \hat{q}[(\mu^R/R) - \mu_R^R]\}\Delta P. \tag{22}$$

A sufficient condition for ΔW and ΔN to have the sign of ΔP is clearly $\mu_P^R \geq \hat{q}\mu_R^R$. It is reasonable, I think, to expect μ_P^R to be non-negative. This means that increases in P, which lower the demand curve facing a fixed-input monopoly, do not raise such a monopoly's optimal price. The sign of μ_R^R seems a priori less clear. The sufficient condition above is satisfied if μ_P^R is non-negative and if, for constant P, the demand curve for the fixed input has constant elasticity at f. If that demand curve is linear, however, it is worth noting that the bracketed expression multiplying q in (22) is negative, so that μ_P^R must then be strictly positive for ΔW and ΔN to have the sign of ΔP. Using Proposition 5 to sign ΔP, we have

Proposition 10: In Case I, if the fixed input monopoly point and the two-part profit maximum are sufficiently close, and if $\mu_P^R \geq \hat{q}\mu_R^R$, then in a move from the former to the latter, ΔP, ΔW, and ΔN all have the sign of $(\bar{q}-\hat{q})$. Such a move is equalizing in both direct and indirect cases.

This analysis provides little support for antitrust hostility toward tying contracts based on monopoly power over fixed inputs. Though the elasticity change condition in Proposition 10 is surely not satisfied in all cases, it surely holds with positive probability in most cases. Moreover, it is not a necessary condition. Even in some situations in which it does not hold, a move from f to π accomplished by a tying contract or equivalent metering device that raises P increases both the number of buyers and aggregate welfare. Such ties tend to be equalizing, so there is no obvious inter-buyer equity problem. Section 7 shows that these properties are also generally present in Cases II and III. Unless a two-part pricing arrangement has serious undesirable long-run effects of some sort or one knows for certain that the sufficient condition in Proposition 10 fails badly, its adoption seems more likely to deserve applause than attack. Note also that tying contracts apparently may result in enough of a drop in the price of the fixed input so that total sales of the variable input rise.

On the other hand, local movements from f to π via sales of the variable input below cost (to permit increases in R) seem socially unattractive on the basis of Proposition 10. Such movements are likely to be equalizing, but this would seem to be overshadowed by reductions in the number of buyers and in total welfare. One might want to take seriously the almost inevitable charges of predation in such cases, even if the monopolist has no predatory intentions.

The analysis of movements from the variable-input monopoly point, point v in Figure 1, to the two-part profit maximum, π, is essentially symmetrical in form to that just presented. The results are a bit weaker, however, reflecting the more complex relation between R and F at π. In terms of Figure 1, we need the slope of the $P^{\pi}(R)$ locus at v. Writing (14b) in terms of the elasticity e^P and differentiating, that slope is given by

$$(dP/dR)^v = [Pe^P N_P - PQe^P_R]/[PQe^P_P + Qe^P(e^P-1)],$$

where e^P_P and e^P_R are the partial derivatives of the elasticity e^P with respect to P and R, respectively, at point v. The denominator is positive by the second-order conditions for a variable-input monopoly.

Substitution of $(dP/dR)^v(\Delta R)$ for (ΔP) in (17) and substitution into (20) yields

$$\Delta W \lesseqqgtr \Delta Q \lesseqqgtr \{Q[Q_R e^P_P - Q_P e^P_R] + Q_P Q_R + e^P Q_P (N_P - Q_R)\}\Delta R. \tag{23}$$

The corresponding expression for ΔN does not simplify usefully. In Case I, Proposition 1 allows (23) to be re-written as

$$\Delta W \lesseqqgtr \Delta Q \lesseqqgtr \{[\hat{q} + (\sigma/Q_R)]e^P_R + (e^P/P) - e^P_P\}\Delta R. \tag{24}$$

As above, it is reasonable to expect e^P_R to be non-negative, while the likely sign of e^P_P seems less clear. In any case, use of Proposition 8 to sign ΔR yields

Proposition 11: In Case I, if the variable-input monopoly point and the two-part profit maximum are sufficiently close, and if $\hat{q}e^P_R \geq e^P_P$, then in a move from the latter to the former, ΔR, ΔW, and ΔQ have the same sign. If $e^P_R \geq 0$, such a move is equalizing in the inverse case.

In the inverse case, Proposition 8 shows that R must increase, and $e^P_R \geq 0$ is sufficient for $P^{\pi}(R)$ to be negatively sloped at v, so that P falls. In the direct case R may rise or fall, so that P may rise or fall also, and the change considered may be equalizing or inequalizing.

The sufficient condition given in Proposition 11 is clearly not necessary. On the same sort of argument made below Proposition 10, Proposition 11 indicates that one can expect tying contracts based on monopoly power over the variable input to increase both the usual welfare measure and total sales of the variable input. If efficiency were the only aim of public policy, this would argue against the current antitrust hostility toward such tying arrangements. The equity implications of ties based on variable-input monopoly power are a bit less clear, however. Even in Case I, the total number of buyers may rise or fall. If e_R^P is non-negative, a move from v to π via a tie that raises R also lowers P and is thus inequalizing in the direct case.

Movements from the variable-input monopoly point to the two-part profit maximum via sales of the fixed input below cost generally lower welfare and sales of the variable input. By Proposition 8, $\bar{q} > \hat{q}$ is necessary for the optimal R to be less than F, so that moving from v to π by selling the fixed input below cost and raising P is most likely equalizing. Thus Proposition 11, like Proposition 10, suggests that tying contracts and related pricing arrangements may be much more socially desirable than persistent sales below cost that aim at achieving the same sort of price discrimination. Proposition 10 makes a stronger case, though.

VII. Local Comparisons: Cases II and III

This Section briefly presents the generalizations of Propositions 9 - 11 to Cases II and III. This requires re-examination of equations

(19), (21), and (23), which hold in all three Cases.

The first term on the right of (19a) and the right-hand side of (19b) are always negative by conditions (11). The second term on the right of (19a) is equal to $(NQ_P\eta)$ in Case II and equal to zero in Case III. Using Proposition 5' to sign the change in P, one can establish

Proposition 9': If the unconstrained welfare maximum and a two-part optimum are sufficiently close, a move from the former to the latter lowers N in Cases II and III, and it also lowers Q in Case III and if $\eta \geq 0$ in Case II. Such a move is equalizing in Case II in both the direct and inverse cases. It is equalizing in Case III in the direct (inverse) case for normal (inferior) inputs.

The third term in brackets on the right of (21) is zero in Case III and $(-N_R\mu_R^R\eta)$ in Case II. Applying Propositions 2, 3, and 5' again, we obtain

Proposition 10': If the fixed-input monopoly point and the two-part profit maximum are sufficiently close, a move from the former to the latter has the following properties. In Case II, ΔP, ΔW and ΔN all have the sign of $(\bar{q} - \hat{q})$ if $\eta \geq 0$ and $\mu_P^R \geq \hat{q}\mu_R^R$, and the move is equalizing in both direct and inverse cases. In Case III, ΔP, ΔW, and ΔN all have the same sign if $\mu_P^R \geq [\hat{q} + (\zeta^R/N_R)]\mu_R^R$, and the move is equalizing in the direct (inverse) case for normal (inferior) inputs.

In Case II, ties may be welfare-reducing for inferior goods with strong income effects even if the elasticity change condition is satisfied.

Finally, consider equation (23). The right-most term in brackets is again zero in Case III, and it equals $(e^PQ_P\eta)$ in Case II. Applying

Propositions 2, 3, and 8' yields

Proposition 11': If the variable-input monopoly point and the two-part profit maximum are sufficiently close, a move from the former to the latter has the following properties. In Case II, ΔR, ΔW, and ΔQ have the same sign if $\eta \leq 0$ and $(Q_P/Q_R)e_R^P \geq e_P^P$; the latter condition alone is sufficient in Case III. If $e_R^P \geq 0$, such a move is equalizing in the inverse case in Case II and in the inverse case for inferior inputs in Case III.

It is interesting to note that normal goods with strong income effects cause trouble here, while inferior goods stand out in the analysis of local movements from point f in Proposition 10'.

Appendix

Define the marginal firm type as a function of Z, R, and P by

$$S(P,Z,\theta^*) = R. \tag{A1}$$

All firms with $\theta \geq \theta^*$ purchase the fixed input. Differentiation of (A1) and use of equations (7) establish

$$\theta^*_R = 1/S_\theta(P,Z,\theta^*) > 0, \tag{A2}$$

$$\theta^*_P = q(P,Z,\theta^*)\theta^*_R > 0, \tag{A3}$$

$$\theta^*_Z = -x(P,Z,\theta^*)\theta^*_R < 0. \tag{A4}$$

The key function Z(P,R) is then defined implicitly by

$$\phi(R,P,Z) = \int_{\theta^*(R,P,Z)}^{1} x(P,Z,\theta)m(\theta)d\theta - D(Z) = 0$$

Differentiation and use of (7) yield

$$\phi_R = -\theta^*_R\, x(\theta^*)\, m(\theta^*) < 0, \tag{A5}$$

$$\phi_P = -\theta^*_R\, q(\theta^*)\, x(\theta^*)\, m(\theta^*) - \alpha\ , \text{ where} \tag{A6}$$

$$\alpha = \int_{\theta^*}^{1} q_Z(P,Z,\theta)m(\theta)d\theta,$$

$$\phi_Z = \theta^*_R\, x(\theta^*)^2\, m(\theta^*) + \delta > 0, \text{ where} \tag{A7}$$

$$\delta = \int_{\theta^*}^{1} x_Z(P,Z,\theta)m(\theta)d\theta - D_Z(Z) > 0. \tag{A7}$$

As is noted in the text below (8), α is positive for normal inputs and negative for inferior inputs. The sign of ϕ_P is thus in general ambiguous for inferior inputs. Note that δ is the derivative of infra-marginal excess supply (holding the set of suppliers fixed) with respect to output price.

Since $Z_R = -\phi_R/\phi_Z > 0$, it follows that an increase in R always raises the output price. Equation (A5) makes it clear that this occurs because such increases exclude marginal suppliers; if Z is unchanged there is no impact on infra-marginal firms. On the other hand, as Panzar and Willig [1978] stress, $Z_P = -\phi_P/\phi_Z$ cannot be signed in general. In the normal case, an increase in P serves both to exclude marginal suppliers and to reduce the supply of infra-marginal firms, thus increasing Z. (Recall that $q_Z = - x_P$ from (7c).) In the inferior case, however, infra-marginal supply is increased, and the net impact on price is ambiguous.

Given Z(R,P), the marginal customer type can be written as a function of R and P only:

$$\hat{\theta}(R,P) = \theta^*[R,P,Z(R,P)]. \tag{A8}$$

Differentiation of (A8) and use of (7) and (A2) - (A4) yield

$$\hat{\theta}_R = \theta^*_R + \theta^*_Z Z_R = \left[\delta/x(\hat{\theta})m(\hat{\theta})\right] Z_R > 0, \tag{A9}$$

$$\hat{\theta}_P = \theta^*_P + \theta^*_Z Z_P = \hat{q}\hat{\theta}_R - \left[\delta/m(\hat{\theta})\right] Z_R. \tag{A10}$$

Since R affects Z only through exclusion of marginal suppliers, the net effect of R must be to lower N, and thus to raise $\hat{\theta}$. Similarly,

$\hat{\theta}_P$ is necessarily positive in the inferior case but may be negative for a normal input. In the latter case, market price is increased as infra-marginal suppliers substitute away from the variable input, and this indirect effect may outweigh the direct exclusionary effect of higher P on marginal firms. If the input is inferior, increases in infra-marginal supply lower Z and thus amplify the direct exclusionary effect.

References

Adams, W. J., and Yellen, J. L. "Commodity Bundling and the Burden of Monopoly." Quarterly Journal of Economics, Vol. 90 (August 1976), pp. 475-98.

Blackstone, E. A. "Restrictive Practices in the Marketing of Electrofax Copying Machines and Supplies." Journal of Industrial Economics, Vol. 23 (March 1975), pp. 189-202.

Blair, R. D., and Kaserman, D. L. "Vertical Integration, Tying, and Anti-trust Policy." American Economic Review, Vol. 68 (June 1978), pp. 397-402.

Bowman, W. S., Jr. "Tying Arrangements and the Leverage Problem." Yale Law Journal, Vol. 67 (November 1957), pp. 19-36.

Buchanen, J. M. "The Theory of Monopolistic Quantity Discounts." Review of Economic Studies, Vol. 20 (1953), pp. 199-208.

Burstein, M. L. "The Economics of Tie-In Sales." Review of Economics and Statistics, Vol 42 (February 1960), pp. 68-73.

Cummings, F. J., and Ruhter, W. E. "The Northern Pacific Case." Journal of Law and Economics, Vol. 22 (October 1979), pp. 329-50.

Diewert, W. E. "Applications of Duality Theory" in M. D. Intriligator and D. A. Kendrick, eds., Frontiers of Quantitative Economics, Vol. II. Amsterdam: North-Holland, 1974.

Feldstein, M. S. "Equity and Efficiency in Public Pricing." Quarterly Journal of Economics, Vol. 86 (May 1972), pp. 175-87.

Gabor, A. "A Note on Block Tariffs." Review of Economic Studies, Vol. 23 (1955), pp. 32-41.

Gould, J. R. "Price Discrimination and Vertical Control: A Note." Journal of Political Economy, Vol. 85 (October 1977), pp. 1063-71.

Kwoka, J. E., Jr. "Output under Second Degree Price Discrimination." Working Paper 21, Bureau of Economics, Federal Trade Commission, October 1979.

Leland, H. E., and Meyer, R. A. "Monopoly Pricing Structures with Imperfect Information." Bell Journal of Economics, Vol. 7 (Autumn 1976), pp. 449-62.

Littlechild, S. C. "Two-Part Tariffs and Consumption Externalities." Bell Journal of Economics, Vol. 6 (Autumn 1975), pp. 661-70.

Mitchell, B. M. "Optimal Pricing of Local Telephone Service." American Economic Review, Vol. 68 (September 1978), pp. 517-37.

Murphy, M. M. "Price Discrimination, Market Separation, and the Multi-Part Tariff." Economic Inquiry, Vol. 15 (October 1977), pp. 587-99.

Ng, Y., and Weisser, W. "Optimal Pricing with a Budget Constraint - The Case of the Two-Part Tariff." Review of Economic Studies, Vol. 41 (July 1974), pp. 337-45.

Oi, W. Y. "A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly." Quarterly Journal of Economics, Vol. 85 (February 1971), pp. 77-96.

Ordover, J. A., and Panzar, J. C. "On the Nonexistence of Pareto Superior Outlay Schedules." Bell Journal of Economics, Vol. 11 (Spring 1980a), pp. 351-54.

Ordover, J. A., and Panzar, J. C. "On the Nonlinear Pricing of Inputs." Discussion Paper 80-14, Department of Economics, New York University, August 1980b.

Panzar, J. C., and Willig, R. D. "On the Comparative Statics of a Competitive Industry with Inframarginal Firms." American Economic Review, Vol. 68 (June 1978), pp. 474-78.

Peterman, J. L. "The International Salt Case." Journal of Law and Economics, Vol. 22 (October 1979), pp. 351-64.

Roberts, K. W. S. "Welfare Considerations of Nonlinear Pricing." Economic Journal, Vol. 89 (March 1979), pp. 66-83.

Scherer, F. M. Industrial Market Structure and Economic Performance, 2nd Ed. Chicago: Rand McNally, 1980.

Schmalensee, R. "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination." American Economic Review, Vol. 71 (1981), forthcoming.

Singer, E. M. Antitrust Economics. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

Spence, A. M. "Multi-Product Quantity-Dependent Prices and Profitability Constraints." Discussion Paper 694, Harvard Institute for Economic Research, April 1979.

Willig, R. D. "Pareto-Superior Nonlinear Outlay Schedules." Bell Journal of Economics, Vol. 9 (Spring 1978), pp. 56-69.

Footnotes

1. Oi [1971] and Blackstone [1975] discuss interesting examples of both explicit and implicit metering. The classic analysis of tying contracts as metering devices is Bowman [1957]. Tying contracts may serve other purposes, of course; see, for instance, Burstein [1960], Singer [1968, chs. 15-17], Gould [1977], Blair and Kaserman [1978], Scherer [1980, pp. 582-4], Cummings and Ruhter [1979], and Peterman [1979].

2. The main works in this literature seem to be Feldstein [1972], Ng and Weisser [1974], Littlechild [1975], Leland and Meyer [1976], and Mitchell [1978]. Gabor [1955] provides an early analysis of profit-maximizing two-part structures.

3. On such tariffs, see Buchanan [1953], Leland and Meyer [1976], Murphy [1977], Willig [1978], Roberts [1979], Spence [1979], Ordover and Panzar [1980b], and the references they cite.

4. The central role of variability here clearly distinguishes the subject of this essay from the fixed-proportion bundling schemes that have been studied by a number of authors, most notably Adams and Yellen [1976].

5. Income elasticities of demand do appear in Feldstein [1972], Littlechild [1975], and Mitchell [1978], but income serves only as a characteristic along which households differ; changes in R do not affect infra-marginal demand. Feldstein's analysis differs sharply from that in this paper in that he treats N as fixed, so that R becomes a lump-sum tax that households cannot avoid. Littlechild errs in neglecting the effect of P on N.

6. One of the important early cases involving tying contracts, for

instance, dealt with a patented machine for attaching buttons to shoes: Heaton Penninsular Button-Fastener Co. v. Eureka Specialty Co., 77 F. 288 [6th Cr., 1896]. One reason for suspecting that interaction through output price might be qualitatively important here is the demonstration by Ordover and Panzar [1980a] that such interaction can undo the main results of Willig [1978] in a situation related to that considered here.

7. Littlechild [1975, p. 666] conjectures but does not prove that $0 < R < F$ could be optimal in Case I.

8. For recent analyses of these pricing strategies, see Kwoka [1979] and Schmalensee [1981], respectively.

9. Mitchell [1978] derives these relations (in elasticity form) after assuming a particular functional form for $S(P,\theta)$.

10. In their analysis of this Case, Ng and Weisser [1974] maximize the integral over $\theta \geq \hat{\theta}$ of $\{I[P,Y(\theta)-R,\theta]m(\theta)\}$ subject to a profitability constraint. They treat $Y(\theta)$ as a control variable, which turns out to be optimally employed to make I_Y an increasing linear function of q_Y. Their formulae cannot be used in the pure profit-maximization case, since changes in P or R do not then invoke the compensating changes in $Y(\theta)$ that Ng and Weisser employ. Moreover, their approach seems less natural in most situations involving public enterprises than that taken in the normative analysis below, where it is assumed (in effect) that $Y(\theta)$ is exogenously fixed and that $I_Y(\theta)$ is constant for all θ.

11. On this theorem, see, for instance, Diewert [1974, Sect. 2].

12. See, for instance, Diewert [1974, Sect. 3]. As Diewert shows, S must be jointly convex in P and Z. Strict convexity in Z is

assumed for simplicity in what follows; this means that firms' marginal cost functions are rising over the relevant range (because $x_Z > 0$) and that firms' input demand functions, holding output price constant, are declining (because $q_p < 0$).

13. Because marginal cost is rising for each firm, an increase in output price must raise the optimal x. Since input prices haven't changed, q increases if and only if the input is normal. See Panzar and Willig [1978] for a clear presentation of these relations. It should be noted that "normal" here corresponds to "strongly normal" there, and "inferior" here corresponds to "strongly inferior" there.

14. See footnote 12, above, on the assumption necessary for σ to be strictly negative. We are operating in a smooth neoclassical world.

15. By "distribution-free", I mean that the social marginal utility of income is taken as unity for all households and firms. Feldstein [1972] presents an alternative framework in which that marginal utility is a given function of income. See also footnote 10, above, on the analysis of Ng and Weisser [1974].

16. Littlechild [1975] proves this for Case I, but he assumes that N does not depend on P. He also modifies his model to allow increases in N to confer positive externalities on infra-marginal customers.

17. If W is maximized subject to a lower bound on Π, and λ is the non-negative associated multiplier, $\omega = \lambda/(1+\lambda)$. It should be clear that the second-order conditions for this constrained maximum problem do <u>not</u> correspond to those for concavity of $\Psi(R,P)$ except when ω is zero or one. On this approach, see Spence [1979].

18. Oi [1971] argues that in Case I non-crossing demand curves ensure $P > V$, and he shows by example that P may be less than V in the inverse case. Littlechild [1975] obtains equation (14) for Case I, except that his neglect of the impact of P on N causes him to have e_p in place of $\bar{e}_p$.

19. Footnote 10, above, describes the differences between Case II and the model of Ng and Weisser [1974].

20. See, for instance, the examples of tying arrangements discussed by Singer [1963, pp. 189-190] and Scherer [1980, pp. 582-4] and the case study presented by Blackstone [1975].

21. Ng and Weisser [1974] deal with income effects without making an assumption of this sort. This stems from their use of the income distribution as a control variable, as footnote 10, above, discusses.

22. By treating the income distribution as a control variable (see footnote 10, above), Ng and Weisser [1974] eliminate this possibility by essentially eliminating income effects.

23. Ordover and Panzar [1980b] have recently investigated necessary and sufficient conditions for the optimality of positive R when $F = 0$ in Case III with identical firms.

24. It is generally assumed that these points are distinct, though it is possible for π to coincide with either f or v.

25. This essay treats tying contracts only as a method of price discrimination; the other functions discussed by the references cited in footnote 1, above, are outside our analysis. Similarly, sales below cost are evaluated only as part of a discriminatory pricing arrangement, though such conduct may in fact have other objectives

such as entry deterrence or predation.

26. Feldstein [1972] is concerned with the linkage between these two sorts of incidence in this context.

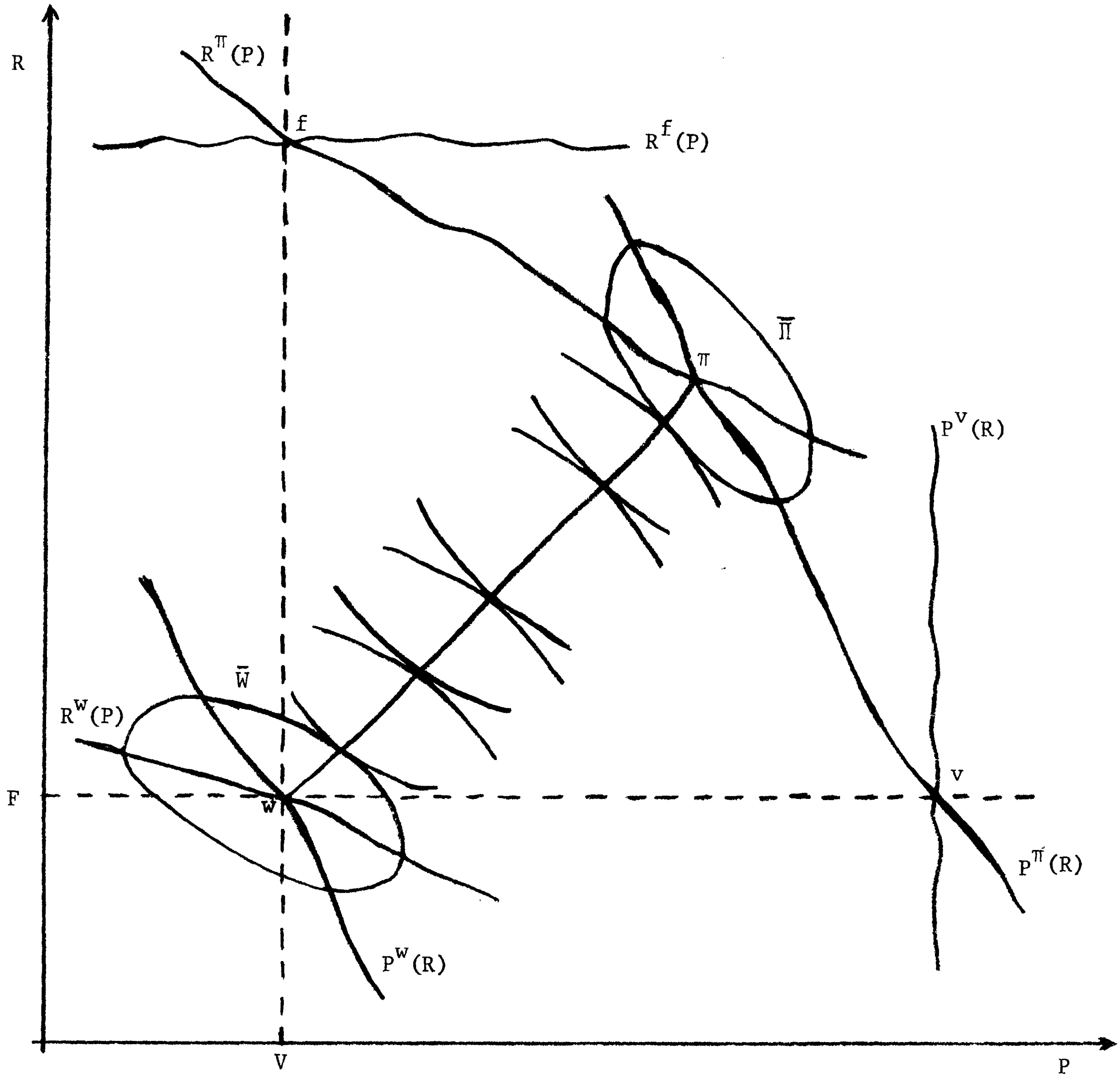

Figure 1

Alternative Equilibria

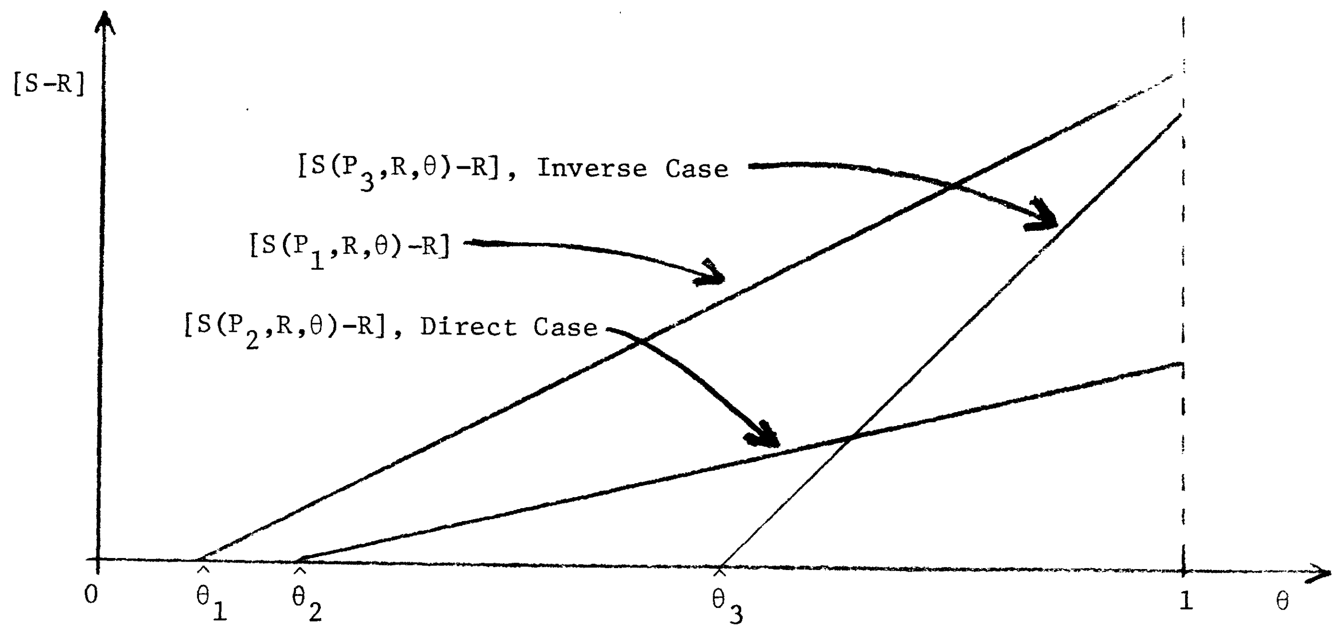

Figure 2. Equalization and Inequalization of Net Surplus: $P_2, P_3 > P_1$