THE WELFARE COSTS OF RATIONING BY WAITING

ROBERT T. DEACON and JON SONSTELIE

With price controls and rationing by waiting, rational consumers increase the quantity bought per purchase. This individually rational response is socially wasteful and the cost of making it is a deadweight loss. This cost plus the value of time spent in queues may exceed the total rent transferred from suppliers to consumers by price controls; i.e., the value of resources spent competing for the rent may exceed the rent itself. This point is illustrated by an empirical application to gasoline price controls. Rent seeking exhausts an estimated 116 percent of the rent transferred.

I. INTRODUCTION

The efficiency of the price system is often illustrated by pointing to the obvious wastes associated with alternative allocation mechanisms, for example the queues that accompany price controls and first come/first served allocations. Although rationing by waiting wastes resources in a dramatically visible fashion, the magnitude of the waste and the factors that determine it seldom have been examined empirically.[1] Moreover, existing theory on the subject is incomplete; unless the first come/first served rule is precisely specified, consumers will compete for available supplies in ways other than waiting. In the model developed below queues cause consumers to compete by increasing amounts bought per trip to the market. The analysis demonstrates that such individually rational adjustments only increase the welfare cost of rationing by waiting, and thus are socially self-defeating.

The theoretical model is illustrated empirically by estimating the welfare costs of a market-wide ceiling on gasoline prices. This application was motivated by the availability of two key pieces of information, the value of time spent in gasoline queues and the effect of waiting lines on amounts bought per purchase. This information was obtained from a survey of motorists conducted during the era of federal gasoline price regulations.[2]

To focus on competition among consumers, one of the inefficiencies that results from price controls is ignored, the supply reduction leading to the

* University of California, Santa Barbara. This research received support from Resources for the Future. Helpful comments were provided by Yoram Barzel, Tom Borcherding, Ross Eckert, Ted Frech, Ron Johnson, Duane Leigh, Dick Sweeney and seminar participants at the University of Oregon and the USC/UCLA Applied Micro-economics Workshops. Constructive criticism from two anonymous referees is gratefully acknowledged.

1. For theoretical treatments of rationing by waiting see Barzel [1974] and Holt and Sherman [1982]. Empirical results have been presented by Frech and Lee [1985] and Alderman [1985]. Lindsay and Feigenbaum [1984] have studied rationing by "waiting lists."

2. See Deacon and Sonstelie [1985].

Economic Inquiry
Vol. XXVII, April 1989, 179–196

loss typically represented by the familiar Marshallian triangle. This issue is avoided by assuming that market supply is fixed.

II. CONSUMER CHOICE WITH NONPRICE RATIONING

In the standard model of demand a consumer chooses only quantities of goods per period. Many goods are purchased through shopping trips to retail outlets, necessitating decisions about the number of trips per period and the quantity purchased per trip, but these decisions are usually suppressed. This is a useful simplification when goods are rationed by market prices because it is the price of a good, not the cost of shopping, that adjusts to balance supply and demand. This is not so useful when goods are rationed by waiting, however, because waiting balances supply and demand by adjusting the cost of a shopping trip. An adequate model of rationing by waiting must therefore extend the standard model, to determine the choice of quantity per purchase and to identify the effect of shopping costs on consumer demand.

Consider a world with two goods, a composite numeraire and a second good that is subject to a price ceiling. For concreteness we refer to the second good as gasoline. The purchase of gasoline requires outlays of time as well as money, even in the absence of rationing by waiting. It simplifies matters to assume that purchases of the numeraire require no expenditure of time.

The following terms are defined for a representative consumer:

x = consumption of numeraire per period,
z = consumption of gasoline per period,
y = quantity of gasoline bought per purchase,
$c(y)$ = cost per period of obtaining y per purchase,
v = value of time,
p = money price of gasoline,
t = time required to make a gasoline purchase, and
$U(x, z)$ = utility.

A consumer who buys y gallons per purchase experiences a time cost of vt/y per gallon. If y could be varied costlessly the rational consumer would purchase a lifetime supply in a single trip and store it for future consumption. This possibility is ruled out by including the per period cost $c(y)$, and assuming $c' > 0$, $c'' > 0$. To avoid ambiguity, $c(y)$ is called an inventory cost.

This cost function results from a choice problem faced by each consumer. A consumer can increase the amount of gasoline bought per purchase in many ways. The most immediate is to decrease the average reserve, to refill the gas tank when it is quarter-full instead of half-full. The cost of this is the increased probability of running out or the inconvenience of an unplanned purchase. Another way to increase y is to expand gas tank capacity, which entails obvious costs. The cost function $c(y)$ merely assumes that the consumer picks the cheapest method to achieve any level of y. Official limits on the amount of gasoline bought per purchase also could enter the cost minimization problem and be represented in the resulting cost function. In fact, rationing by waiting often induces limits on quantity per purchase, and the following analysis shows why this may occur. Such limits give the cost function an infinite slope at the quantity limit, if enforcement is perfect. If enforcement is not perfect, as is likely, the cost minimization problem should reflect all actions a consumer might take to circumvent the limit, such as search for retail outlets where enforcement is lax. These activities may cause the slope of the cost function to become very steep near the official limit, but this does not alter the analysis or conclusions that follow.

The term t includes both time needed to process a transaction and any time spent waiting in a queue. The time needed to process a transaction is of little interest here and accordingly is held fixed. To ease the exposition and to direct attention to the phenomenon of primary importance, t is simply called "waiting time" in the following discussion.

All welfare effects are measured with the compensating variation, which applies only if consumers receive lump-sum transfers sufficient to maintain constant utility. Accordingly, compensated demand functions are used to represent consumer responses.[3] These functions result from the following expenditure minimization problem:

$$\underset{x,y,z}{\text{minimize}} \qquad x + \theta z + c(y) \tag{1}$$

$$\text{subject to} \qquad U(x, z) = U^0$$

$$\text{where} \qquad \theta = p + vt/y$$

and U^0 is the consumer's initial utility level. The objective function is the consumers' dollar outlay for x and z, plus the value of time spent waiting to purchase z, plus the inventory cost associated with buying y units per purchase. The term θ is the full price of gasoline in the sense of Becker [1965].[4] The solution to this problem gives the choice variables as functions of p and t, denoted $\hat{x}(p, t)$, $\hat{y}(p, t)$, and $\hat{z}(p, t)$, as well as the expenditure minimizing full price, $\theta(p, t) = p + vt/\hat{y}(p, t)$. The utility level is suppressed as an argument in these functions because it remains fixed throughout the analysis.

3. This is the approach recommended by Diamond and McFadden [1974].

4. The full price of z will decline as z increases, so that a locus of constant expenditures on x and z will be convex to the origin. Thus, the sufficient second-order condition for expenditure minimization requires that the U^0 indifference curve be more convex than this locus.

The full price ties the functions $\hat{x}$ and $\hat{z}$ to the usual notion of compensated demand. The first-order conditions for problem (1) reduce to three equations. The first two are the utility constraint and the equality of full price with the marginal rate of substitution between x and z. These equations implicitly define x and z as functions of the full price and the utility level. They thus define Hicksian demand functions for x and z, denoted $x^*(\theta)$ and $z^*(\theta)$. These Hicksian demands are related to the functions $\hat{x}$ and $\hat{z}$ by

$$\hat{x}(p, t) = x^*[\hat{\theta}(p, t)] \tag{2}$$

$$\hat{z}(p, t) = z^*[\hat{\theta}(p, t)].$$

The third first-order condition from problem (1) is

$$vtz/y^2 = c'(y) \tag{3}$$

which states that the marginal benefit y yields in reduced waiting time just balances the marginal cost of larger purchases. This equation implicitly defines y as a function of z and t, $\tilde{y}(z, t)$, and gives the value of y that minimizes the transactions cost of z units of gasoline, $vtz/y + c(y)$. The relationship between $\tilde{y}$ and $\hat{y}$ is

$$\hat{y}(p, t) = \tilde{y}[\hat{z}(p, t), t]. \tag{4}$$

The function $\tilde{y}$ is convenient in what follows.

To evaluate the equilibrium waiting time one needs two comparative statics results from the expenditure minimization problem. To obtain the first of these, differentiate $\tilde{y}$ partially with respect to t:

$$(\partial\tilde{y}/\partial t)\,(t/y) = c'/(yc'' + 2c') = (\eta + 2)^{-1}. \tag{5}$$

$\eta = yc''/c' \geq 0$ is the elasticity of marginal inventory cost (c') with respect to y. According to (5), if consumption of z per period is held constant, an increase in waiting time increases the amount bought per purchase. The responsiveness of y to t depends on η, the elasticity of c' with respect to y. The larger is η the more nearly vertical is the marginal inventory cost schedule and the smaller is the resulting adjustment in y. The limiting case occurs where η is infinite, so that marginal cost is vertical and y is fixed.

To obtain the second result, differentiate $\hat{z}(p, t)$ in (2) with respect to t and substitute from (5). After some rearrangement the following expression emerges:[5]

5. The derivation of (6) is available from the authors.

$$\partial\hat{z}/\partial t = (\partial\hat{z}/\partial p)\,(v/y)\,(\eta + 1)/(\eta + 2). \tag{6}$$

To interpret this notice that if η were infinite, and hence y were fixed, (6) would imply that $\partial\hat{z}/\partial t = (\partial\hat{z}/\partial p)\,(v/y)$. With y fixed, increasing the waiting time by dt is equivalent to raising the money price by $(v/y)dt$. In the more general case where η is finite, increasing t will cause consumers to make larger purchases. The increase in y mitigates the increase in full price and hence the change in z. In general, then, $|\partial\hat{z}/\partial t| < |\partial\hat{z}/\partial p|\,(v/y)$, and the difference between the two sides of this inequality is larger the smaller is η.

III. EQUILIBRIUM WAITING TIME

Now introduce n consumers, indexed by the subscript $i = 1, \ldots, n$. In equilibrium their compensated demand functions will satisfy the market clearing condition

$$\sum_{i=1}^{n} \hat{z}_i\,(p, t) = \bar{z} \tag{7}$$

where $\bar{z}$ represents the fixed supply of gasoline. Equation (7) establishes a functional relationship between p and t. In the absence of a price ceiling, t is simply the technologically fixed amount of time needed to process a transaction. Given this fixed value of t there is a unique value of p that equates supply and demand. If p is controlled at a level below equilibrium, a queue develops and t rises to clear the market for z. Thus, for any p less than or equal to the uncontrolled equilibrium, there is a unique t that satisfies equation (7). Denote this equilibrium relationship

$$t = t^*(p). \tag{8}$$

Equations (6) and (7) can now be used to examine this function.

Finding the derivative of (8) will reveal the factors that determine the equilibrium waiting time for a given ceiling price. To proceed, differentiate the market clearing condition (7) with respect to p and substitute from (6); this yields

$$(dt^*/dp) = -\sum (\partial\hat{z}_i/\partial p)/[\sum (\partial\hat{z}_i/\partial p)\,(v_i/y_i)\,(\eta_i + 1)/(\eta_i + 2)]. \tag{9}$$

To most easily interpret this, imagine all consumers are identical. Dropping subscripts, (9) simplifies to

$$(dt^*/dp) = -(y/v)\,(\eta + 2)/(\eta + 1). \tag{10}$$

Fixing p below its market clearing level causes a relatively long waiting time if the value of time is small or the quantity bought per purchase is large. This is not surprising. The term in (10) involving η, however, indicates that the waiting time will be long if consumers find it easy to increase quantity per purchase. Intuitively, queues cause consumers to economize on waiting costs by increasing y, the amount bought per purchase. As y rises the full price falls, however, and excess demand appears. This excess demand increases the waiting time and the full price until equilibrium is restored. With identical individuals the same full price clears the market with or without the adjustment in y. The full price is invariant because increases in the amount bought per purchase shift neither the demand function nor the supply of the product. Increases in y do, however, result in larger inventory costs, a point which is explored later.[6]

The relationship between t and p with nonidentical consumers is more complicated, but carries exactly the same intuition. Inspection of equation (9) confirms that $|dt^*/dp|$ is decreasing in v_i and η_i, and increasing in y_i. A given ceiling price will therefore result in a relatively long waiting time if the value of time is low, the quantity per purchase is large, or the quantity per purchase is highly responsive to waiting time.

IV. A GENERAL MEASURE OF CONSUMER WELFARE COST

The total social loss due to a price control is the sum of the rent lost by suppliers and the net welfare cost, positive or negative, experienced by consumers. The following focuses on consumers only and uses the compensating variation as a welfare cost measure. When a price ceiling is imposed and queues develop, the appropriate compensating variation is the total compensation, expressed in units of numeraire, required to maintain constant utility for all consumers. Considering a typical consumer, and using the compensated demand functions (x^* and z^*) and the expenditure minimizing full price (θ_i) the minimum expenditure on x and z required for utility U^0 is

$$M_i(\hat{\theta}) = x^*(\hat{\theta}) + \hat{\theta} z^*(\hat{\theta}). \tag{11}$$

6. Increases in y raise the *waiting time per purchase* (t) and lower the number of *purchases per period* (z/y), leaving unchanged the time spent *waiting per period* (tz/y). The result in the text can also be shown by stating the market-clearing condition in a slightly different way. With identical consumers and fixed supply the market-clearing full price, here denoted $\bar{\theta}$, cannot be affected by a price ceiling. Hence, $p + vt/y = \bar{\theta}$ for all values of p below the uncontrolled equilibrium, which implies $t = (\bar{\theta} - p)\,(y/v)$. Given any p, increases in y cause proportional increases in t.

To this must be added the inventory cost $c(\hat{y})$ to obtain the total outlay needed to maintain constant utility. The compensation required to maintain fixed utility after imposition of a price ceiling is therefore

$$W = M(\theta^1) + c(y^1) - M(\theta^0) - c(y^0), \tag{12}$$

where the superscripts 0 and 1 denote expenditure minimizing values of variables in the pre- and post-control regimes.

Recalling that $\partial M/\partial\theta = z*(\theta)$, the compensating variation is

$$W = \int_{\theta^0}^{\theta^1} z*(\theta)d\theta + c(y^1) - c(y^0). \tag{13}$$

The integral expression in (13) is the area to the left of the compensated demand schedule between prices θ^0 and θ^1. Since the price ceiling normally will cause full prices to fall for some consumers and rise for others, this area will be negative for some consumers and positive for others. The second term in (13) is the net cost from increasing the amount bought per purchase. The consumer benefits from the price ceiling if W is negative and loses if W is positive.

To obtain a more useful expression for W, integrate (13) by parts:

$$W = \theta^1 z^1 - \theta^0 z^0 - \int_{z^0}^{z^1} z0^{z1}\, \theta*(z)dz + c(y^1) - c(y^0), \tag{14}$$

where $\theta*(z)$ is the inverse compensated demand function. Recalling the definition of θ, rewrite (14) as

$$W = (p^1 z^1 - p^0 z^0) + [\,(vt^1 z^1/y^1) - (vt^0 z^0/y^0)] - \int_{z^0}^{z^1} \theta*(z)dz + [c(y^1) - c(y^0)]. \tag{15}$$

This four-part expression represents the cost to a representative consumer of a price ceiling. The sum of these expressions across all n consumers is the aggregate consumer loss. When summed in this manner, the first term reduces to $(p^1 - p^0)\bar{z}$ because supply is fixed; that is, $\sum z_i^0 = \sum z_i^1 = \bar{z}$. The amount $(p^1 - p^0)\bar{z}$ is the transfer from suppliers to consumers in the form of lower money prices. While it is a benefit to consumers, it is exactly offset by losses to suppliers and thus represents neither a net gain nor loss to society.

The second term in (15) is the value of time spent waiting in queues. Because there is no offsetting gain to anyone from this use of resources, it represents a deadweight loss.

The third term in (15) reflects the welfare consequences of reallocating z among consumers. The pre-control allocation is taken to be Pareto efficient, and imposing a price ceiling will, in general, cause this allocation to change. Accordingly, a reallocation of z and x among consumers could make all better off. When summed over all consumers, the third term is the amount by which aggregate compensation could be reduced if that reallocation were made.[7]

Finally, consumers who attempt to economize on waiting times by making larger purchases incur inventory costs. The fourth term in (15) shows the magnitude of this cost increase for an individual consumer.

Given the pre-control price and waiting time, p^0 and t^0, the welfare cost expressions in (13) and (15) are functions of p^1 and t^1. Since the post-control equilibrium waiting time can be expressed as a function of the ceiling price, it follows that the *equilibrium* welfare cost can be written as a function of the ceiling price alone, $W = W(p^1)$.

Before examining the form of $W(p^1)$ in a fully general context, consider once again a world of identical consumers. In that world all must face the same full price both before and after the price ceiling is introduced. To maintain the same full price, the increase in the value of time spent waiting must exactly offset the decrease in money price, implying that the first two terms of equation (15) cancel. The third term in equation (15), which results from the misallocation of z among consumers, must be zero if all consumers are identical. The entire consumer loss is therefore captured in the fourth term, the cost of increasing the amount bought per purchase. This reinforces the self-defeating nature of these adjustments. In the absence of such behavior the consumer loss function $W(p^1)$ would be identically zero for all ceiling prices.

With nonidentical consumers more complex outcomes are possible, as Figure 1 illustrates. The compensated demand schedules of two consumers are D_1 and D_2. To simplify the diagram it has been assumed that $t = 0$ in the uncontrolled equilibrium; thus $\theta_1^0 = \theta_2^0 = p^0$. The price ceiling specifies a money price of p^1, and the post-control equilibrium full prices for persons 1 and 2 are denoted θ_1^1 and θ_2^1.

7. In the present context a Pareto-efficient allocation can be characterized by the first-order conditions for the problem

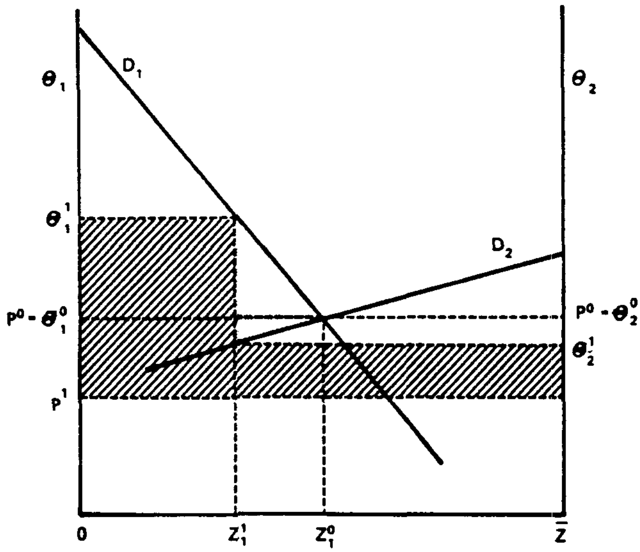
$$\max_{(x_i,\, y_i,\, z_i,\, t)} \quad W[U_1(x_1, z_1), \ldots, U_n(x_n, z_n)]$$

$$\text{subject to} \quad \sum [x_i + (z_i t v_i / y_i) + c_i(y_i)] = \bar{x}$$

$$\sum z_i = \bar{z}, \text{ and } t \geqq \bar{t},$$

where $\bar{x}$ and $\bar{z}$ are aggregate supplies of x and z, and $\bar{t}$ is the technologically determined minimum time required to make a purchase. The first-order conditions require that the terms $(\partial U_i/\partial z_i)/(\partial U_i/\partial x_i) - (v_i \bar{t}/y_i)$ be identical for all $i = 1, \ldots n$. That is, consumers' marginal rates of substitution, net of *minimum* transaction costs, must be equalized. A price ceiling causes a waiting time in excess of $\bar{t}$ and hence violates this condition.

FIGURE 1



The first component of the welfare loss expression in (15) is the rent transferred from suppliers. It is negative, representing a gain to consumers, and appears in Figure 1 as a rectangle with base $0\bar{z}$ and height $p^0 - p^1$. Offsetting this gain is the cost associated with time spent waiting, indicated by the diagonally lined area. An additional efficiency loss stems from the misallocation of z among consumers, depicted by the shaded triangle. The costs associated with increasing amounts bought per purchase are not incorporated in Figure 1.

A case where a price ceiling causes consumers to lose in aggregate was deliberately chosen for Figure 1, because this outcome seems less intuitive than the alternative. Notice that rationing by waiting reallocates the fixed supply away from the consumer whose demand is inelastic and toward the consumer whose demand is elastic. Thus the total expenditure of time and money required to maintain pre-control levels of utility is increased unambiguously, and aggregate consumer welfare is reduced. Stated differently, the full price has been raised to the consumer with inelastic demand and lowered to the consumer with elastic demand. This is the same general pattern of price discrimination that would be practiced by a monopolist who sought to increase total consumer spending for a fixed amount of product. Naturally, if the direction of reallocation in Figure 1 were reversed, aggregate consumer welfare would be increased.[8]

8. Frech and Lee [1985] discuss the effects of different elasticities on the welfare cost of price controls.

To determine the welfare effect of a price ceiling when consumers are not identical, differentiate (13) with respect to p^1 and sum across consumers:

$$d(\sum w_i)/dp^1 = \sum [z_i^* (\theta_i^1)(d\theta_i^1/dp^1) + c_i'(y_i^1)(dy_i^1/dp^1)]. \qquad (16)$$

Using equation (3), this reduces to

$$d(\sum w_i)/dp^1 = \sum z_i^* (\theta_i^1) + (1/t^1)(dt^1/dp^1)[\sum z_i^* (\theta_i^1) v_i t^1/y_i^1]. \qquad (17)$$

The first sum in (17) is fixed because supply is fixed, so attention can be focused on the second sum. Recalling that $dt^1/dp^1 < 0$, a marginal reduction in the money price can either raise or lower aggregate consumer welfare; that is, $d(\sum w_i)/dp^1$ may be either positive or negative. The sum in brackets will be relatively large, and hence consumer losses relatively likely, if there is a positive correlation between individual consumption (z_i^*) and the value of time per gallon purchased (v_i/y_i^1). The two-person example in Figure 1 incorporates just such as positive correlation.

Finally, consumers tend to lose from a price ceiling if they are flexible in increasing amounts bought per purchase. This is implied by the finding in equation (9) that flexibility in adjusting y_i causes $|dt^1/dp^1|$ to be large. For another perspective on this, recall that a consumer who increases y_i incurs an inventory cost in return for the benefit of a reduced waiting cost, and at the margin these terms are equal. By increasing y_i, however, the individual also causes the waiting time to rise for all other consumers. Since this increase in t is not offset by benefits to anyone, it represents a deadweight loss for consumers as a group.

V. AN EMPIRICAL APPLICATION

The waiting time and welfare cost functions, equations (8) and (15), are estimated for a price ceiling on gasoline. Most of the information required for these estimates emerged from a natural experiment with rationing by waiting that occurred during 1980. Due to an anomaly in federal price regulations a small number of Chevron stations in California were required to reduce their prices well below the market. Their supplies were rationed on a first come/first served basis, and long queues formed. Most service stations were not so constrained, and sold gasoline at market clearing prices without queues. In Deacon and Sonstelie [1985] results from a survey of patrons of high-price and low-price stations were reported. Choice of stations was examined in a probit equation that estimated the value of time spent waiting (v_i) as a function of income, employment status, and other variables. The relationship between quantity bought per purchase (y_i) and expected waiting time was estimated in a second equation. The surveyed motorists here are treated as the universe of gasoline consumers in order to simulate the effects of imposing a market-wide price ceiling.

The Equilibrium Waiting Time Function

Limitations on available data force the adoption of two simplifying assumptions. First, individual compensated demand functions are locally approximated by linear functions with identical price slopes,

$$z_i^1 = z_i^0 + b(\theta_i^1 - \theta_i^0) \qquad i = 1 \ldots n \tag{18}$$

where $b < 0$. Recalling that $\sum z_i^1 = \sum z_i^0$, sum these terms across consumers to obtain

$$n^{-1} \sum \theta_i^0 = n^{-1} \sum \theta_i^1$$

or

$$\mathrm{E}(\theta_i^0) = \mathrm{E}(\theta_i^1) \tag{19}$$

where $\mathrm{E}(\cdot)$ is expectation operator. Thus the average full price paid by consumers is unaffected by the price ceiling. Inserting the definition of θ into (19) yields

$$t^1 = [p^0 - p^1 + t^0\mathrm{E}(v_i/y_i^0)]/\mathrm{E}(v_i/y_i^1) \tag{20}$$

where it should be noted that y_i^1 depends on t^1.

The second assumption is that the inventory cost function takes the constant elasticity form

$$c_i(y_i) = h_i y_i^{(\eta+1)} \tag{21}$$

where h_i is a parameter specific to the individual. Given this functional form, the elasticity of marginal inventory cost with respect to y_i is η, which coincides with earlier notation. From (5) the elasticity of y_i with respect to t is $(\eta+2)^{-1}$, here denoted γ; this implies

$$y_i^1/y_i^0 = (t^1/t^0)^\gamma. \tag{22}$$

A given increase in the waiting time causes the same proportionate increase in gallons bought per purchase for all consumers. Substituting (22) into (20) yields a simplified expression for the equilibrium waiting time,

$$t^1 = \left\{[(p^0 - p^1)/\mathrm{E}(v_i/y_i^0)] + t^0\right\}(t^1/t^0)^\gamma. \tag{23}$$

Controlling for differences in sizes of gasoline tanks, the consumers surveyed bought 53 percent more gasoline per purchase if they patronized the low-price stations. This implies a point estimate of .1532 for γ.[9]

Patrons of the high-price stations often purchased when their tanks still held substantial quantities of gasoline. Those who bought from the low-price stations, however, typically would not enter a queue unless their tanks were near empty. We assumed that adjusting the reserve in this fashion is the only way consumers changed purchase size. This seems appropriate since the pricing anomaly lasted only a few months, not long enough to allow sensible changes in tank capacity. Accordingly, capacity was taken as given and the econometric equation for purchase size controlled for differences in sizes of gasoline tanks. If permanent price controls were instituted consumers would adjust more fully and along more margins, and the response elasticity of purchase size would presumably be larger than the estimate used here.

Given t^0 and an estimate of $\mathrm{E}(v_i/y_i^0)$ from the survey, the equilibrium waiting times reported in Table I were computed for various values of p^0-p^1.[10] Two estimated waiting times are reported for each price difference. The values labelled "y_i variable" follow directly from equation (23) and allow for adjustments in quantity per purchase. The relationship between t and p is nonlinear in this case; successive \$.05 reductions in the ceiling price cause more than proportionate increases in the queue. If y_i is held fixed at y_i^0, much shorter waiting times result. Fixing y_i also makes the waiting time linear in p^0-p^1); each additional \$.05 reduction in the ceiling price lengthens the waiting time by 2.15 minutes.

The Welfare Cost Function

The welfare cost formula in equation (15) requires information on θ_i^0, θ_i^1, z_i^0, z_i^1, the form of the compensated demand functions, and the inventory

9. Average waiting times were 0.89 minutes and 14.6 minutes, respectively, at the unregulated and price controlled stations. Consumers who chose to wait at a controlled station bought about 53 percent more gasoline, holding tank capacity constant, than those who chose an unregulated station. Using equation (22), this implies that $\gamma = .1532$.

10. The expected waiting time at the unregulated stations (t^0) was 0.89 minutes, or .0149 hours. The terms y_i^0 and v_i, measured as gallons per purchase and dollars per hour, were estimated from equations for choice of station and quantity per purchase in Deacon and Sonstelie [1985]. For the sample of survey respondents, $\mathrm{E}(v_i/y_i^0)$ equalled 1.3981.

TABLE I
Equilibrium Waiting Times

Price Difference p^0-p^1, (\$/gal.)	Waiting Time per Purchase (minutes) y_i variable[a]	y_i constant[b]
.05	3.79	3.04
.10	7.13	5.19
.15	10.73	7.33
.20	14.53	9.48
.25	18.49	11.62

[a] Computed from $y_i^1/y_i^0 = (t^1/t^0)^\gamma$ where $\gamma = .1532$.

[b] Computed assuming $y_i^1 = y_i^0$.

cost. The data needed to estimate θ_i^0 and θ_i^1 were discussed earlier. To identify z_i^0 and z_i^1 for each consumer we rely on the assumed linearity of compensated demand functions, an external estimate of the price slope of these demand functions, and survey information on quantities consumed per period.[11]

The inventory cost function specified in equation (21) implies $c_i' = (\eta+1)c_i/y_i$, or

$$c_i(y_i^1) - c_i(y_i^0) = (\eta+1)^{-1}\,[c_i'(y_i^1)y_i^1 - c_i'(y_i^0)y_i^0]. \tag{24}$$

Substituting equation (3) into this yields a form amenable to estimation,

$$c_i(y_i^1) - c_i(y_i^0) = (\eta+1)^{-1}[(v_i t^1 z_i^1/y_i^1) - (v_i t^0 z_i^0/y_i^0)]. \tag{25}$$

To identify $(\eta+1)$ in this equation, recall that $\gamma = (\eta+2)^{-1}$ and thus $(\eta+1)^{-1} = \gamma/(1-\gamma)$.

11. Given z_i^0 for a consumer who patronized an unregulated station, z_i^1 is estimated from equation (18) using information on b and $(\theta_i^1 - \theta_i^0)$. A similar procedure is used to estimate z_i^0 for those who bought from a price-controlled station. To estimate b in equation (18) we rely on an external estimate of the aggregate uncompensated price elasticity of demand for gasoline. Writing individual demands as $z_i = a_i + b\theta_i = a_i + bp + bv_it/y_i$, it follows that the aggregate demand may be expressed as $n\mathrm{E}(z_i) = n\mathrm{E}(a_i) + nbp + n\mathrm{E}(v_it/y_i)$. With demand estimated as a function of money price, the price elasticity of aggregate demand would be $\delta = bp/\mathrm{E}(z_i)$. Hence, we estimate b as $\delta\mathrm{E}(z_i)/p$ using $\delta = -.20$, the median of short-run gasoline price elasticity estimates reported in Morlan et al. [1981]. The money price, p, was placed at \$1.25 per gallon, and $\mathrm{E}(z_i)$ was computed as gasoline sales per licensed driver per day in California. See California Energy Commission [1981], and California Department of Finance [1982].

TABLE II
Equilibrium Welfare Costs
(dollars per consumer per day)

Difference p^0-p^1 ($/gal.)	y_i variable: Transfer (1)	Waiting Cost (2)	Mis-allocation (3)	Inventory Cost (4)	Total (5)	y_i constant: Total (6)
.05	−.1345	.1318	.0001	.0238	.0213	−.0023
.10	−.2689	.2635	.0003	.0477	.0425	−.0052
.15	−.4034	.3950	.0006	.0715	.0636	−.0078
.20	−.5379	.5263	.0010	.0952	.0846	−.0104
.25	−.6724	.6574	.0014	.1189	.1054	−.0135

Note: (1) Transfer $= \frac{1}{n}(p^1-p^0)\bar{z}$

(2) Waiting cost $= \frac{1}{n}\sum_i \left(\frac{v_i t^1 z_i^1}{y_i^1} - \frac{v_i t^0 z_i^0}{y_i^0} \right)$

(3) Misallocation $= \frac{1}{n}\sum_i \left[\int_{z_i^0}^{z_i^1} \theta_i^*(z_i, U_i^0)\, dz_i \right]$

(4) Inventory cost $= \frac{1}{n}\sum_i \left\{ c_i(y_i^1) - c_i(y_i^0) \right\}$

(5) Total $=$ (1) + (2) + (3) + (4)

(6) Total welfare cost per person, assuming $y_i^0 = y_i^1$, $i = 1, \ldots n$.

Table II presents the welfare cost estimates. Positive entries indicate compensation that must be paid to consumers to maintain pre-control utilities, and hence represent welfare losses. Two sets of estimates are presented for each price difference. The first group, labelled "y_i variable," takes into account changes in amounts bought per purchase, the resulting effects on waiting times, and welfare costs. The total welfare costs reported in column (5) are all positive and indicate net consumer losses from price controls. For example, if the price were reduced by \$.25 per gallon, consumers on average would gain \$.6724 per day because of the lower price, lose \$.6574 per day in waiting costs, lose \$.0014 per day because of the misallocation of gasoline among themselves, and lose \$.1189 per day from the cost of increasing gallons per purchase. The total loss averages \$.1054 per day, which amounts to nearly \$.04 per gallon, since the average consumption is 2.69 gallons per day. On net, then, a \$.25 reduction in the money price actually raises the total cost of gasoline by \$.04 per gallon.

All four components of the welfare cost function are approximately proportional to the price difference p^0-p^1. The rent transfer is necessarily proportional since supply is taken as fixed. Waiting cost, inventory cost, and total welfare cost rise at rates that decrease slightly as the price difference increases. Only the misallocation component grows at an increasing rate.

The entries in column (6) of Table II show that consumers would experience small gains from a price ceiling were it not for adjustments in purchases. Our unpublished results indicate that preventing such adjustments eliminates the inventory cost item, with little effect on other welfare cost components.[12]

Individual variations in values of time, amounts purchased, and consumption per period cause these welfare effects to vary among consumers. Table III presents welfare effects for nine consumer groups, together with individual and family incomes to permit distributional comparisons. All estimates pertain to a ceiling price \$.25 per gallon below market. Estimated welfare effects range from a gain of \$.16 per day for part-time workers to a loss of \$.45 per day for the high income fully-employed. Overall, these estimates exhibit mild progressivity. The three groups who experience gains have lower than average incomes, while two of the three groups that incur the largest losses have the highest reported incomes. The pattern of progression is not monotonic, however. Among the four groups reporting family incomes of \$15,962 to \$18,697, welfare effects range from a gain of \$.16 per day to a loss of \$.11 per day.[13]

VI. CONCLUSION

When evaluating the empirical results, certain aspects of the price control episode examined should be kept in mind. Because it lasted only a few months, observed adjustments in gallons purchased and associated variations in full price are best regarded as short-run in nature. In a long-run setting, consumer responses to queues might be more dramatic, e.g., increasing

12. These results are insensitive to the aggregate price elasticity used to estimate individual demand slopes. An elasticity of -.50 was used to represent long-run responses in place of the -.20 elasticity used in Table II. This substitution approximately doubled the "misallocation" costs shown in Table II, but had no appreciable effect on the other components.

13. The welfare cost estimates in Tables II and III differ from those in Deacon and Sonstelie [1985]. Estimates in the former paper represent welfare effects for consumers who *chose* to wait in line in a situation where only a few service stations were price-controlled. In that context, only those who stood to gain waited in line. The present paper estimates the welfare effects on all consumers of a price ceiling imposed market-wide.

TABLE III
Distributional Effects of Rationing by Waiting
($p^0 - p^1$ = \$.25/gallon)

Occupational Category	Welfare Cost ($/day)	Individual Income ($/year)	Family Income ($/year)
Part-time Workers	−.1617	$ 10,375	$18,625
Students	.0477	$ 2,647	$ 2,647
Housewives	−.1093	$ 5,769	$15,962
Unemployed	−.1580	$13,839	$17,277
Fully Employed:[a]			
Income ≤ $10,000	.2849	$ 7,500	$11,222
$10,000 < income ≤ $20,000	.1094	$14,713	$18,697
$20,000 < income ≤ $30,000	.0639	$24,730	$31,841
$30,000 < income ≤ $40,000	.2268	$34,351	$35,833
$40,000 < income	.4494	$52,500	$57,917

Note: Figures in each column are averages for individuals in various groups.
[a] Income ranges for the fully employed refer to individual income.

gasoline capacity. This would, of course, lengthen the waiting time and increase welfare losses.

In the incident studied, all consumers could obtain gasoline on known terms from regular supply sources. There were no uncertainties about the availability of gasoline, as occurred during the brief price control crises of 1974 and 1979. Those uncertainties elicited an entirely different response; consumers increased their average gasoline reserve by frequently "topping off" their tanks. This is a rational response if one fears that additional gas may be unavailable at any price. It is not relevant to the situation studied, however, where a certain waiting time clears the market and consumers know that they can always buy gasoline by enduring that wait.

Our results confirm the widely held dictum that price controls, with supplies allocated on a first come/first served basis, result in waiting costs that dissipate at least part of the rent transferred from suppliers. Competition among consumers to "be first" results in queues, and thus is self-defeating. If the quantity bought per purchase can be varied, consumers will compete along a second margin by increasing quantity per purchase. This, too, is self-defeating because it lengthens the waiting time and wastes the cost of making such adjustments.

There is no apparent reason why competition among consumers should be confined to only two margins, waiting and adjusting amounts bought per purchase. The forms competition might take seem to be limited only by the institutional details of the allocation rule used. With gasoline waiting lines, for example, those with high values of time have an incentive to hire others with low values of time to wait in line on their behalf, e.g. teenagers filling the family car, employees buying gasoline for the boss, and so forth. Such arrangements presumably involve transactions costs, but it may be privately rational to incur these costs in order to lower the value of time spent waiting, and thereby to obtain a lower full price. Any reduction in full prices, however, would cause excess demand to appear and the waiting time to rise to restore equilibrium. Thus this entire process is self-defeating in much the same way that increases in y were found to be. Moreover, the costs of arranging such transactions represent wastes in addition to those associated with waiting and adjusting purchases.[14] It is tempting to speculate that these examples illustrate a general principle, that welfare losses rise with increases in the number of margins for consumer response to rationing by waiting.

Finally, the analysis explains the quantity limits that often accompany rationing by waiting. If quantities per purchase were frozen at levels that would prevail without price controls, the deadweight loss of inventory cost adjustments would be avoided. It is unlikely that a perfect scheme of quantity limits could be instituted, however. Quantity per purchase without controls would surely vary widely across consumers, necessitating an equally wide variety of quantity limits when prices are controlled. The sort of universal limit often observed is at best a crude approximation to a perfect scheme. Furthermore, any scheme of quantity limits would be difficult to enforce, tempting consumers to engage in costly actions to circumvent them. To the extent that this avoidance behavior was successful, the result would be socially self-defeating in the same manner as adjustments in the average gasoline reserve.

14. Holt and Sherman [1982] reach a similar conclusion in their analysis of "waiting-line auctions."

REFERENCES

Alderman, Harold. "Allocation of Goods Through Nonprice Mechanisms: Evidence on Distribution by Willingness to Wait." Photocopy Washington, D.C.: International Food Policy Research Institute, 1985.

Barzel, Yoram. "A Theory of Rationing by Waiting." *Journal of Law and Economics*, April 1974, 73–96.

Becker, G. S. "A Theory of the Allocation of Time." *Economic Journal*, September 1965, 493–517.

California Department of Finance. *California Statistical Abstract: 1982*. Sacramento, 1983.

California Energy Commission. *Quarterly Fuel and Energy Summary*. Vol. 6, No. 4, Fourth Quarter 1980, Sacramento, 1981.

Deacon, Robert T. and Jon Sonstelie. "Rationing by Waiting and The Value of Time: Results from a Natural Experiment." *Journal of Political Economy*, August 1985, 627–47.

Diamond, Peter A. and Daniel L. McFadden. "Some Uses of the Expenditure Function in Public Finance." *Journal of Public Economics*, February 1974, 3–21.

Frech, H. E. and William C. Lee. "The Welfare Loss of Rationing by Queuing Across Uses: Theory and Estimates from the U.S. Gasoline Crises." *Quarterly Journal of Economics*, February 1987, 97–108.

Holt, Charles A., and Roger Sherman. "Waiting-Line Auctions." *Journal of Political Economy*, April 1982, 280–94.

Lindsay, Cotton M. and Bernard Feigenbaum. "Rationing by Waiting Lists." *American Economic Review*, June 1984, 404–17.

Morlan, Terry H., Daniel H. Skelly, and Arnold P. Reznek. "Price Elasticities of Demand for Motor Gasoline and Other Petroleum Products." U.S. Department of Energy, Energy Information Administration. Washington: U.S. Government Printing Office, 1981.