# Price Distortion and Shortage Deformation, or What Happened to the Soap?

By MARTIN L. WEITZMAN*

*The model of this paper generalizes the classical theory of consumer behavior to the more general case of prices that are not necessarily market-clearing. Suppose that, in addition to the money cost, some sort of search, waiting, or other quasi-fixed "effort-cost" is needed to obtain goods. The presence of this quasi-fixed cost element will trigger an inventory policy. A shortage equilibrium occurs when effort costs are such that, in the corresponding inventory policy, the flow of desired consumption does not exceed the available supply flow. Stock hoarding, a critical phenomenon in the economics of shortage, emerges as a natural component of this model. A complete characterization of a stationary shortage equilibrium is given. Comparative statics and welfare analysis are performed. The dynamic transition between steady states is analyzed to give insight into the mechanics of how shortages develop.* (*JEL* D50)

It is known, in a general way, that price distortions lead to shortages, queues, searching, hoarding, and so forth. Yet it seems fair to say that the exact mechanism integrating each main element of a "shortage syndrome," especially the stockpiling phenomenon, has not been clearly articulated.[1] The main aim of this paper is to provide a usable model of shortages by appropriately generalizing the classical theory of consumer behavior to a situation where prices are not necessarily market-clearing. The model essentially consists of an equilibrium approach that combines inventory theory with demand theory.

A particularly vivid illustration of the phenomenon I have in mind is illustrated by recent Soviet experience in the consumer-goods market. Consider Soviet soap as a metaphorical example. The example is metaphorical because, while the shortage phenomenon in the consumer-goods market I seek to describe is quite general, the particular commodity most illustrative of the general phenomenon can vary.

Throughout most of 1989 (at the time this paper was written), there was virtually no soap available on the shelves of Soviet stores. When officials in charge of planning were asked about this problem, they acted embarrassed and annoyed. In newspaper articles and television talk shows, they explained repeatedly that production this year was actually up 10 percent over last year, which was itself 4 percent higher than the previous year. Furthermore, not only was production accelerated somewhat as this embarrassing shortage became evident, but more than $8 million of valuable foreign exchange was spent on buying soap abroad. Finally, they pointed out that statistics of

*Department of Economics, Harvard University, Cambridge, MA 02138. This research was supported by a grant from the National Science Foundation.

[1]Some models dealing with somewhat different aspects of this phenomenon are described in Janos Kornai and Jorgen Weibull (1978), Victor Polterovich (1983), Dale Stahl and Michael Alexeev (1985), Kent Osband (1989), and the references cited in these works. The case typically treated has the waiting-effort cost proportional to the amount of the good bought. This is equivalent to the assumption that people are limited to buying one small unit at a time and must wait in line anew for each small purchase that is made. (It is then a straightforward exercise to generalize to an "as if" market equilibrium where the "as if" price is the money price plus the appropriately normalized disutility of waiting-effort price.) I much prefer the opposite assumption: after waiting in line for a sufficiently long time, or happening upon the good, the customer can effectively buy as much as he wants. I think this is a more realistic assumption of the two extremes; additionally it leads naturally to an analysis of the key inventory-stock aspect of the problem.

per capita soap consumption show the Soviet Union to be not very far behind the advanced Western capitalist countries. All of this sounds quite believable; and on the whole Soviet citizens do not seem to be going unwashed.

The above story can be repeated for any number of commodities. Soviet leaders frequently characterize the economy as being in a "crisis situation" ("krizisnoye polozheniye"); but what, exactly, is the crisis? What is the appropriate model of causality leading from budget deficits and a monetary overhang to disorders in the consumer-goods markets? What is happening? What should be done?

To such questions no clear answers emerge. Some cite breakdowns in the distribution system. (Railroads seem particularly to be accused.) Others blame a "hoarding psychology" that causes panic buying and is somehow related to the deficit and monetary overhang. Theft by workers, sabotage, and speculation by cooperatives are also candidates. The officials seem unified only on promising increased production to meet the shortage and on calling for formation of committees to investigate formally the problem.

Some informal investigation reveals an interesting, if perhaps not unexpected, fact. Although few official figures are available, observations, conversations, and anecdotes suggest strongly that Soviet people are hoarding soap and other commodities in massive amounts. Significant parts of bathrooms, closets, hallways, and other areas have been given over to storage.

My aim is to model carefully the general process of "hoarding psychology," which is a fairly widespread occurrence in shortage situations, even if it is not typically so extreme as the above case. I believe the model has potential applications to a wide variety of situations where prices are stuck at "wrong" values for whatever reason. Thus, some conclusions may have relevance for malfunctioning markets in capitalist economies and may even help to understand certain features of "fixed-price" macroeconomics. It will be shown that hoarding psychology can be given a quite rational economic interpretation and can be coherently analyzed within the appropriately extended framework of standard economic theory.

## I. The Model

Suppose there are $n$ goods in the economy, denoted $i = 1, 2, \ldots, n$. For the sake of argument, all consumers are assumed to be identical. (Allowing consumers to be different would not affect the existence or general form of an equilibrium, but it would render less sharp the characterization of its properties.)

Each consumer has the same utility function of the form

$$U(\mathbf{d}) - V(e) - W(s). \tag{1}$$

In the above formula, $\mathbf{d} = (d_i)$ is the usual consumption $n$-vector. The variable $e_i$ stands for the amount of "effort" required to obtain good $i$ each time that it is obtained. Conceptually it is perhaps easiest to think of $e_i$ as the time wait in line needed to buy good $i$. (There is a separate line for each good, and after waiting the required time,[2] the consumer can buy as much as desired.) However, $e_i$ could more generally represent search effort of any sort expended to obtain the commodity. On the most abstract level, $e_i$ is interpreted as the degree of difficulty in obtaining good $i$. If good $i$ is obtained at frequency $f_i$, then the total effort (per unit time) expended on obtaining all goods[3] is

$$e \equiv \sum f_i e_i. \tag{2}$$

While it is conceptually easiest to think of effort $e_i$ as deterministic, there is no prob-

[2]Note that $e_i$ represents the time wait in line, not the length of the line. If $m$ people, each of whom stocks up amount $s$, are waiting in line and the total flow of goods into the store is $d$, then the time wait in line is $e = ms/d$.

[3]By writing disutility as a function of total effort, $V(e)$, I am implicitly assuming that the sum of a large number of fixed costs incurred at different times can, in effect, be smoothed.

lem with an interpretation that makes each $e_i$ an independently distributed random variable so long as it is small relative to $n$. Then by the law of large numbers, $e$ itself will be (almost) deterministic, equal to a weighted sum of the form (2), where $e_i$ is now interpreted as the *expected* effort that must be spent on obtaining good $i$.

The variable $s_i$ stands for the stock of good $i$ that is purchased and must be stored when the good is obtained. The coefficient $h_i$ represents the opportunity cost per unit of good $i$ carried (per unit time). The magnitude of $h_i$ would reflect such things as the opportunity cost of storage space $i$ takes up (on shelves, in refrigerators, in warehouses, or whenever applicable), shrinkage, the cost of guarding, interest forgone, the inconvenience of hoarding, and so forth. Total storage cost is then

$$s \equiv \sum h_i s_i. \tag{3}$$

The representative consumer's utility function is assumed to be of the additively separable form (1). The first element $U(\mathbf{d})$ is just the traditional utility function of classical consumer theory, having all the usual properties. The function $V(e)$ represents the disutility of effort, while $W(s)$ is the disutility of storage. The underlying assumption in (1) of independence would appear not to be terribly restrictive, and perhaps even reasonable, in the present context. More general formulations could be treated but with some loss in crispness of results. It is assumed that $U(\mathbf{d})$ is smoothly concave, while $V(e)$ and $W(s)$ are smoothly convex.

The representative consumer faces fixed nominal prices $\mathbf{p} = (p_i)$ and is endowed with nominal money income $I$. Thus, the usual budget constraint

$$\mathbf{pd} \leq I \tag{4}$$

holds.

Unlike the classical setup, however, in the situation here prices are not necessarily market-clearing.[4] Quantities available per capita are $\mathbf{q} = (q_i)$. Any feasible consumer demand must satisfy the additional constraints

$$\mathbf{d} \leqq \mathbf{q}. \tag{5}$$

The supplementary constraints (5) distinguish the present model from classical consumer theory. In the classical case, in effect $q_i = \infty$ for all $i$, or else $\mathbf{p}$ represents equilibrium prices that just exactly make demands $\mathbf{d}$ equal to supplies $\mathbf{q}$. For the classical case, in effect there are no explicit constraints on consumer purchases other than the overall budget constraint (4). Here, the interesting case is when constraint (5) "bites" for some goods, representing inadequate supply at the fixed prices, presumably arising ultimately from production limitations.

The present formulation is sufficiently rich to cover a number of special situations of interest. For example, the price might be artificially repressed on only one or a few goods, which then become "deficit" compared to the bulk of commodities, which are market-clearing; or there could be a general deficit of commodities, meaning the prices of most goods are artificially low relative to incomes and availability. Also covered is the case in which the same good is available cheaply in limited amounts at state stores and simultaneously at market-clearing prices in private stores. Yet another situation covered with only slight modification of the present framework is that in which a given fixed vector of goods is to be allocated, so that consumers end up with the same final allocation of goods in any case, but for some given subset of (deficit) goods prices are frozen at below market-clearing levels, while for the remaining (available) goods, prices

[4] It is beyond the scope of this paper to speculate on why certain prices may be set at below market-clearing levels in certain circumstances. Suffice it to note here that the practice is extremely widespread, and there is an extensive literature on many aspects of it. Governments are fearful of raising prices once they have become established, often with good reason because people do not like price increases. The present paper is limited to analyzing the effects of too low prices without delving deeply into the issue of why they are too low in the first place.

move freely to their competitive levels, which just clear the fixed supplies taking account of consumer income. Since $U(\mathbf{d})$ represents the familiar utility function of goods consumed, it automatically embodies the usual relations of complementarity, substitutability, diminishing returns, and whatever else might be considered relevant to the situation at hand.

## II. Coefficient of Price Distortion

In what follows, it will be useful to have a quantitative measure of the degree to which values and prices are distorted in the economy under consideration. To that end, consider the following mathematical programming problem:

$$\max[U(\mathbf{d})] \tag{6}$$

subject to

$$\mathbf{d} \leqq \mathbf{q} \tag{7}$$

$$\mathbf{pd} \leq I. \tag{8}$$

The constrained optimization (6)–(8) is a classical resource-allocation problem. In the present context, it can be interpreted as a second-best problem in optimal rationing. Let the solution be

$$\mathbf{d} = \mathbf{d}^*. \tag{9}$$

Let $\lambda$ be the shadow price of constraint (8). Without significant loss of generality, suppose that the marginal utility of income is positive or that $\lambda > 0$. [The marginal utility of an extra ruble is greater than zero, which means that constraint (7) is not so tight in every component that no goods are available to be bought on the margin. This would be guaranteed in theory if, for example, some of the goods were available in unlimited supply, or if prices on some of the goods were market clearing.] Then, it is not difficult to see that

$$\lambda = \min_i \left[\frac{U_i}{p_i}\right] \tag{10}$$

where

$$U_i \equiv \left.\frac{\partial U}{\partial d_i}\right|_{\mathbf{d}=\mathbf{d}^*}. \tag{11}$$

Necessary and sufficient conditions for an optimum are then

$$U_i > \lambda p_i \to d_i^* = q_i. \tag{12}$$

While $\lambda p_i$ represents the nominal price of an extra unit of good $i$ (normalized so that the marginal utility of income is 1), $U_i$ measures the actual value of an additional unit of good $i$, or what people would actually be willing to pay. In the present context, it is then natural to define the coefficient of value distortion or price distortion of good $i$ (weighted by the amount of the good) as

$$\delta_i \equiv (U_i - \lambda p_i) d_i^*. \tag{13}$$

The coefficient $\delta_i$ measures the difference between the actual value of good $i$ that might be consumed and the nominal price value, normalized in terms of utility. Therefore, $\delta_i$ is a measure of the degree of disequilibrium deviation from market-clearing of good $i$. If the logic of (13) is accepted, the appropriate measure of overall value or price distortion in the economy becomes

$$\delta \equiv \sum \delta_i. \tag{14}$$

The coefficient $\delta$ is a measure of the degree of overall disequilibrium in the economy.

## III. The Basic Problem

The question now arises as to how the goods are actually distributed in the model economy, given the existence of constraints (4) and (5). As a point of departure, suppose that a state of chronic shortages exists in an orderly stationary equilibrium. That is, every consumer knows he must expend effort $e_i$ to obtain good $i$ and plans the appropriate inventory policy. After, in effect, paying the fixed waiting-time cost of $e_i$, the consumer chooses to buy and stock the amount $s_i$ and run it down at the consumption flow rate $d_i$. This pattern is repeated at frequency

$$(15) \qquad f_i = \frac{d_i}{s_i}.$$

Furthermore, economy-wide this repetitive behavior is self-reinforcing.

Of course this description of shortage behavior as a regular steady-state equilibrium with recurrent sawtooth-patterned inventories is an abstraction. Shortage phenomena can be notoriously erratic. Nevertheless, treating a shortage economy as if it were in a stationary equilibrium yields important quantitative insights. Furthermore, it is a necessary first step to any analysis of dynamics. Actually, the methodological issues connected with modeling a shortage equilibrium do not seem fundamentally different from those involved in modeling a nonshortage equilibrium.

Consider the problem facing the typical consumer. Effort levels $\{e_i\} \geq 0$ are taken as given.[5] Consumption flow levels $\{d_i\}$ and inventory stocks $\{s_i\}$ should be chosen to:

$$(16) \quad \underset{\{d_i, s_i\} \geq 0}{\text{maximize}} \left[ U(\{d_i\}) - V\left( \sum e_i \left[ \frac{d_i}{s_i} \right] \right) - W(\Sigma h_i s_i) \right]$$

subject to

$$(17) \qquad \sum p_i d_i \leq I.$$

In what follows, I assume that the first-order necessary conditions for characterizing an optimum to the above problem [(16) and (17)] are also sufficient.[6]

In the economy being modeled, the appropriate equilibrium concept is the following.

DEFINITION: *A stationary shortage equilibrium is a set of $\{e_i, d_i, s_i\}$ satisfying the following three conditions*:

$$(18) \qquad \{d_i, s_i\} \textit{ solves problem } (16)\text{–}(17) \textit{ for given } \{e_i\}$$

$$(19) \qquad d_i \leq q_i \textit{ for all } i$$

$$(20) \qquad \textit{if } d_i < q_i, \textit{ then } e_i = 0.$$

The reader should feel satisfied, upon reflection, that conditions (18)–(20) represent the correct generalization of classical consumer equilibrium theory to the present context.[7]

It is not difficult to generalize the definition of a stationary shortage equilibrium to a situation with many nonidentical consumers having different utility functions and different incomes, nor is it difficult to prove existence using methods similar to those employed in the present paper. This route is not pursued in detail here simply because, as with most truly general equilibrium formulations, it is impossible to characterize sharply the properties of a solution without placing more structure on the model.

The following theorem completely characterizes a stationary shortage equilibrium.

[5] In equilibrium, $\{e_i\}$ will be like an implicit price that equilibrates the system, but each individual consumer will view $\{e_i\}$ as exogenously given.

[6] A variety of conditions would guarantee this result. Essentially, the inventory part is introducing an economy of scale into an otherwise convex problem. So long as the inventory nonconvexity effect can be bounded [e.g., by assuming that $\{e_i\}$ is small relative to the curvature of $U(\cdot)$], the necessary first-order conditions for (16)–(17) would remain sufficient. In this paper, I will not go further into the essentially technical and messy aspect of insuring that necessary conditions are sufficient for the given problem. I am indebted to Victor M. Polterovich for pointing out to me that some assumption needs to be made in order to presume that the necessary first-order conditions for a maximum of (16)–(17) are also sufficient in the present context.

[7] There is an implicit assumption behind definition (18)–(20) that at any instant in time consumers hold uniformly distributed stocks of good $i$, ranging from $s_i$ to 0, and they arrive uniformly to market just as their stock of the good is depleted. This assumption could presumably be justified as the outcome of a dynamic optimizing process in which consumers satisfying the overtaking criterion always choose the line with the shortest wait and thereby force all waiting lines for the same good to have equal length in equilibrium.

THEOREM 1: *The unique stationary shortage equilibrium is the solution of the following equations.*[8]

$$d_i = d_i^* \tag{21}$$

$$s_i = \frac{\delta_i}{wh_i} \tag{22}$$

$$e_i = \frac{\delta_i^2}{vwh_i d_i^*} \tag{23}$$

*where*

$$w \equiv W'(s) \tag{24}$$

*is the marginal disutility of storage evaluated at* $s$ $(=\Sigma h_i s_i)$ *satisfying*

$$sW'(s) = \delta \tag{25}$$

*while*

$$v \equiv V'(e) \tag{26}$$

*is the marginal disutility of effort evaluated at* $e$ $(=\Sigma f_i e_i)$ *satisfying*

$$eV'(e) = \delta. \tag{27}$$

PROOF:

In the optimization problem (16)–(17) that defines condition (18), let $\mu \geq 0$ be the shadow price multiplier for inequality (17). The corresponding necessary and, by assumption, sufficient first-order conditions for any $\mathbf{d}, \mathbf{s} \geq 0$ satisfying (17) to be the unique maximizer of (16) are, for all $i$,

$$U_i - V'\left(\frac{e_i}{s_i}\right) = \mu p_i \tag{28}$$

$$V'\left(\frac{e_i d_i}{s_i^2}\right) = W'(h_i). \tag{29}$$

[Equation (29), rearranged, is the famous square-root law of inventory theory.]

The rest of the proof is essentially by inspection. Assignments (21)–(27) and the supplementary assignment

$$\mu = \lambda \tag{30}$$

are proposed as the solution of (18)–(20). Using conditions (4), (5), (12), and (13), it is straightforward to verify that the proposed solution, (21)–(27) and (30), does indeed satisfy (28), (29), (19), and (20).

With the shortage equilibrium expressed in the simple closed form (21)–(27), it is easy to perform comparative-statics exercises.

Note from (22) that $s_i$ is proportional to $\delta_i$, while from (23) $e_i$ is proportional to $\delta_i^2$. Thus, small shortages show themselves primarily in increased stock hoarding, with just very small increases in search activity. On the other hand, large shortages result in large inventories and very large waiting lines or search times. These observations suggest how shortages might evolve or devolve.

While from (23) effort per purchase is proportional to the square of the coefficient of price distortion, total effort is not. This is because, as waiting lines increase, the consumer reacts by buying bigger bundles less frequently. The total effort per unit time spent on obtaining good $i$ is, from (15) and (21)–(23),

$$f_i e_i = \delta_i / v. \tag{31}$$

Substituting (31) and (22) into (13) yields, respectively,

$$U_i = \lambda p_i + v\frac{f_i e_i}{d_i} \tag{32}$$

$$U_i = \lambda p_i + w\frac{h_i s_i}{d_i}. \tag{33}$$

Thus, the difference between the intrinsic value of a good and its nominal price is made up by the effort expended per unit of the good to obtain it [equation (32)] and also by the cost expended per unit of the good to store it [equation (33)]. In shortage

[8]Conditions (21)–(27) are presented in the given sequence and form to facilitate their economic interpretation. From the strictly mathematical standpoint of presenting an algorithm that uniquely solves (18)–(20), it is preferable to think of the following order: first (21) defines $\{d_i\}$, then (25) defines $s$ and (27) defines $e$, then (24) defines $w$ and (26) defines $v$, then (22) defines $\{s_i\}$ and (23) defines $\{e_i\}$. Equations (25) and (27) have unique solutions in $s$ and $e$, from the convexity of $V(e)$ and $W(s)$.

equilibrium, the consumer ultimately gets what goods [equation (21)] he would have gotten under the optimal rationing scheme (6)–(8). However, from (32) and (33), it is the cost of nonproductive search and storage activity that raises the "as if" equilibrium price from $\lambda p_i$ to $U_i$, the hypothetical price that would have to be paid to clear the "as if" competitive market.

In the simplified world of this model, it is easy to analyze the welfare loss of a distorted price system. As a reference point, I will take the constrained second-best optimal-rationing solution (9) of problem (6)–(8). Such an allocation maximizes utility subject to budget constraint (4) and goods-availability constraint (5). In some sense, the utility difference between it and the shortage equilibrium, to be denoted $L$, represents the waste of search and storage activity. There is a simple expression linking $L$ with the coefficient of price distortion $\delta$.

THEOREM 2: *The welfare loss of a stationary shortage equilibrium compared with an optimal rationing scheme is*

$$L = \delta\left(\frac{1}{a} + \frac{1}{b}\right) \tag{34}$$

*where $a$ is the elasticity of disutility of effort*

$$a \equiv \frac{v}{V}e \tag{35}$$

*while $b$ is the elasticity of disutility of storage*

$$b \equiv \frac{w}{W}s. \tag{36}$$

PROOF:

From (21), the allocation of goods is the same in (18)–(20) as in (6)–(8). Therefore, from (1) the difference in utility attained is

$$L = V(e) + W(s). \tag{37}$$

Making use of definitions (25), (27), (35), and (36), expression (37) can be rewritten as (34).

Note that shortage losses are first-order in the appropriately normalized measure of economy-wide price distortion. This is because, unlike the standard second-order deadweight loss (from tax theory) of distorted prices in markets that clear, here malformed prices cause real shortage deformations akin to rent-seeking activity.[9] Furthermore, the social cost of these shortage deformations enters twice; that is, distorted prices here do double damage. First, they create nonproductive search activity at a social cost of $\delta/a$, which has no other function than to allocate the artificially underpriced goods. Second, price distortions result in goods being tied up by buyers at social cost $\delta/b$ in socially unnecessary inventories held throughout the system.[10]

Somewhat paradoxically, the higher the elasticity of disutility of effort or storage, the less damage is done by a faulty price system. This is because search or storage activities do not play a directly productive role here. Their only purpose is to allocate artificially underpriced goods. Since buyers can only end up consuming on average what is being made available on average, the goods will end up getting distributed the same way no matter what are the $V(e)$ or $W(s)$ functions. When people suffer greater incremental pain from additional search or storage activity, they will actually end up with less total disutility, because they will simply avoid waiting in lines or stocking up their closets. Hence, the paradoxical conclusion is that greater potential pain is less actual pain.

[9]For an analysis of rent-seeking activity, see James Buchanan et al. (1980) and the references cited therein.

[10]The reader may wonder why distorted prices do double damage in this model. For concreteness, take the linear case $a = b = 1$. Then, (34) becomes $L = 2\delta$. Yet in a standard model where the waiting-effort cost is proportional to the amount of the good bought, the total loss of consumer surplus would be equal to the price subsidy times the quantity purchased, or $\delta$. Where, then, does the extra term of $\delta$ come from? In a standard model, there is no need to hold inventories. In the present square-root inventory model, the marginal cost of carrying inventory is always equal to half the average cost. Since equilibrium essentially requires that the marginal cost of holding inventory be equal to the amount of price subsidy, it follows that the average cost of holding inventory must equal twice the price subsidy. I am indebted to Paul R. Milgrom for providing this intuitive explanation of why $L = 2\delta$ when $a = b = 1$.

There is another seeming paradox here, which makes an important point. In an environment of widespread shortage due to a malfunctioning price system, ordinary notions of "real income" lose their meaning. Actually, if there is sufficient price distortion, higher income can, other things being equal, mean lower welfare. If enough goods are in shortage, more "real income" can translate primarily into longer lines and greater hoarding rather than increased consumption per se.

Other things (including prices) being held equal, when nominal income is raised there are two effects. More money is available to spend on nondeficit goods $(i|U_i = \lambda p_i)$, which increases welfare. This is the usual sense in which higher real income is better. However, in a shortage situation there is also a detrimental side, which might be called the "money-overhang effect." As income is raised, the marginal utility of income $(\lambda)$ declines, which increases the difference between value and price $(U_i - \lambda p_i)$ for deficit commodities, which in turn leads to greater search $(e_i)$ and storage $(s_i)$ effort and raises the welfare loss $L$. This money-overhang effect is more pronounced as the deficit commodities constitute a greater fraction of all goods. In the extreme case, the money-overhang effect can be so severe that more income can actually result in lower welfare.

There is a slightly different way of constructing and interpreting the model that makes the same basic point. Take as given a fixed vector of goods $\mathbf{q}$ to be allocated. Consumers end up with the same final allocation of goods $\mathbf{d} = \mathbf{q}$ no matter what. For a given subset of (deficit) goods, prices are frozen at below market-clearing levels. For the remaining (available) goods, prices move freely to their competitive levels, which just clear the fixed supplies, taking account of consumer income. As money income increases in this setup, the consumers will be made unambiguously worse off. Price distortion on the available goods will always be zero. However, for the deficit commodities, the coefficient of price distortion will increase with income as the marginal utility of income declines. For the same final allocation of goods, lower fixed prices or higher nominal income decreases consumer welfare.

Thus, standard measures of "real income" may be a quite unreliable indicator of welfare in situations of economic shortage. It is then not difficult to understand the temptation to eliminate search and storage costs by imposing rationing, even though such measures introduce problems of their own.[11]

In the model as presented, the production or supply side is taken as more or less given. If the supply side were adversely affected by shortages, either because people are so busy finding and keeping goods that they have less time to work or because it is difficult to maintain morale on the job when pay buys so little, the ingredients are present for a vicious circle.[12] An increased money overhang decreases welfare, which then feeds back to lower production, which depresses welfare further. In such a world, income-increasing tendencies that exacerbate a money overhang can be very dangerous indeed.

This point can be illustrated rather simply as follows. Suppose that a unit of labor-time input produces a unit of homogeneous output. Each person is endowed with one fixed unit of labor-time, which can be divided continuously between producing output or engaging in non-directly-productive search and storage activity. The delivery price of the single good (or, in an alternative interpretation, the wage) is fixed by the government at $\omega$, whereas the sales price of the good is fixed by the government at $p < \omega$. Thus, the subsidy of cost over price per unit of the good is $\gamma \equiv \omega - p$. Suppose each person must pay a fixed per capita tax of $\theta > 0$ to the government. Let $y$ be the amount of the good produced per capita. (The remaining time $1 - y$ is spent on search and storage activity.) Then per capita dis-

[11]Throughout this paper, black-market activities are ignored. It could be argued that pressure toward black markets might increase with the degree of price distortion.

[12]Indeed, the Soviet economy is showing signs of this vicious circle.

posable income is $\omega y - \theta$, which in equilibrium must equal the value of goods purchased $py$. Rewriting this latter relation gives $y = \theta / \gamma$. Other things being equal, the greater the price subsidy $\gamma$, the greater is the incentive to engage in nonproductive search and storage activity, and the lower is per capita production (and consumption) of actual goods and services. Note too that other things being equal, a higher per capita tax (within the relevant range) actually raises productive output, because it lessens the monetary overhang caused by price subsidization. It should be readily apparent that if this kind of economy gets caught in any kind of spiral between lower standards of living and increased money-wage subsidization, production can implode.

For the sake of a sharp characterization, the analysis has proceeded as if all consumers are identical. A shortage equilibrium will typically exist in the more general case when consumers are different, but it will be more complicated to analyze. Consumers with less income or a lower value of time will end up waiting in line and storing the subsidized products, while people with higher incomes or a greater value of time will tend to seek out the higher-priced goods with shorter lines.[13]

The model of shortage equilibrium presented here presupposes a regular, stationary, well-behaved process. Actually, shortage phenomena frequently display an erratic, nonstationary aspect. To that extent, the analysis presented here probably understates the social loss of price distortions.

## IV. Dynamics

The analysis so far has been purely static. A stationary shortage equilibrium is postulated. Everyone expects this steady state to persist forever. Consumer behavior, which is contingent upon this stationary equilibrium, reinforces it. As has been shown, the analysis of steady-state shortage equilibrium is eminently tractable. Some suggestive insights emerge about changes in behavior across steady states, but there has been no dynamic analysis as such.

In a model such as this, where inventories are playing a central role, dynamic transitions display an important property not visible from examining steady-state behavior alone. With a change between steady states comes a change in average inventory stocks, which in turn forces a corresponding alteration of consumption flows during the interim. Adjustments toward higher stocks of equilibrium inventories, without a change in supply, must involve an accumulation at the expense of consumption, a kind of "inventory squeeze" of consumption. This in turn leads to an overshooting property in the waiting time of queues, so that short-run welfare temporarily falls below its long-run equilibrium level. The purpose of the present section is to explore this theme in some detail. It turns out that a formal analysis of dynamics for the general case is an extremely formidable task. All that I am able to do here is to provide a suggestive example of the dynamics of one simple case. I believe this example captures well the main elements that are involved in the dynamic properties of an "inventory squeeze." I also believe that the basic features of the example must generalize. However, disequilibrium dynamics is a notoriously difficult area about which rather little is known except that a lot of different things typically can happen. I do not want to claim that all adjustment processes will be as smooth as the example I will describe in detail.

Suppose, then, that the economy starts out in some steady-state shortage equilibrium. This old steady-state equilibrium had been expected to continue indefinitely. Then, suddenly and without warning, some underlying parameters change, which increase the degree of shortage in the economy. To be specific, suppose $\delta_i$ increases to $\delta_i'$, but the underlying supply flow of goods does not change. The easiest thing to envisage is an increase of nominal income that exacerbates the degree of monetary over-

[13] It is even theoretically possible that low-income, low-value-of-time people could be made better off by price subsidies, although this possibility is excluded in a representative consumer world.

hang in the economy. Suppose the new situation is naively expected to last forever.

Using (21)–(23), it is easy to analyze the differences between the two steady states. With $\delta_i' > \delta_i$, the new stationary equilibrium will feature more search effort and larger inventories. Yet this cannot be the entire story of the transition. When the underlying production structure has not changed, consumption will be the same (equal to production) in both steady states, or $d_i' = d_i$. If consumption in the new stationary equilibrium is to be the same as the old, because what has changed is not the underlying production structure but the price of good $i$ relative to income, the economy must go through a wrenching dislocation in making a transition between steady states. There is no way that inventory stocks can be built up from an average per capita level of $s_i/2$ to a higher average per capita level of $s_i'/2$ and have consumption flows remain uninterrupted throughout. With consumption levels the same in beginning and final steady states, but inventory levels higher in the final state, there must be a transition period during which consumption is curtailed to allow stocks to accumulate. In turn, the only allocation mechanism available to force lower consumption is increased waiting time or search costs.

A change from a lower to a higher state of equilibrium shortage involves a change from a lower to a higher level of search time. The greater required effort of the new steady state is bad news enough. The effort level during the transition period is even more awful news because, to induce consumers to curtail consumption and permit inventory stocks to accumulate, effort must "overshoot" its new long-run equilibrium and then only gradually decline to the new steady-state level, which is merely worse than the old.

Suppose the change occurred at time $t = 0$. At time $t < 0$, everyone was expecting the old shortage equilibrium to last forever. Then, at time $t = 0$, $\delta_i$ unexpectedly increased to $\delta_i'$. This new condition, too, is expected to be permanent. The typical consumer, when he now goes to wait in line to make his purchases and restock his inventories at time $t \geq 0$, will attempt to adjust. Projecting the same effort level to obtain the good but at a higher value relative to price, the consumer will initially attempt to buy and to consume more of the good. Since the flow of desired purchases would then exceed the flow of available goods and since price is fixed (by assumption), the length of the queue must increase to throttle back desired purchases to available supply. Even should waiting lines have increased enough to trigger the new steady-state inventory policy, in which desired consumption flow is equal to available supply flow, that is not enough. In the new equilibrium, each consumer wants to hold more inventories and make purchases less often because the ratio of search effort to nominal price is higher than before. That means that every buyer coming to market in the time immediately after $t = 0$ wants to buy more than before in order to stock up to the new inventory level. Then, waiting times must be even higher than the new equilibrium level, to keep potential purchasers away. The effort level to obtain the good in the transition period must be so high that potential buyers are forced to delay purchases by cutting back on consumption and waiting it out until the lines come down to the new equilibrium level.

In Figure 1 are depicted typical "before," "during," and "after" patterns of effort, inventory, and consumption. At least this is the pattern of the example I wish to present. Before the shock at $t = 0$, inventory stocks for a representative consumer display the usual sawtooth pattern with periodicity

$$T \equiv \frac{s}{d}. \tag{38}$$

(For notational convenience, the subscript $i$ will henceforth be dropped whenever its use is superfluous from the context.) After adjustment, the new steady-state inventory sawtooth has periodicity

$$T' \equiv \frac{s'}{d}. \tag{39}$$

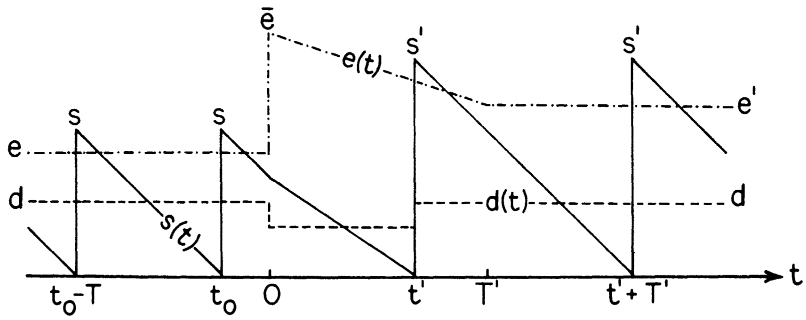


FIGURE 1. TYPICAL TRAJECTORIES OF EFFORT, CONSUMPTION, AND INVENTORY

I will exhibit one relatively simple dynamic transition path that lasts exactly $T'$ time units. At time $t < 0$, the economy is in the old steady-state equilibrium. During time $0 \leq t \leq T'$, the economy is in a state of transition. For times $t > T'$, the economy is in its new steady-state equilibrium. The transition itself is a dynamic equilibrium because everyone believes it will happen this way, and their subsequent optimizing behavior makes it actually come about. There may be other consistent transition paths, but I believe they must demonstrate the same basic "inventory squeeze" aspect of the path I will describe, only in more complicated form, because average inventory stocks cannot be built up without somehow squeezing consumption flows in the process. I am not bothered by limiting myself to the relatively simple example of a transition described here because: a) it is already complicated enough, b) it illustrates well the basic principles involved, and c) there is little purpose to misplaced generality here when the model itself and the other aspects of dynamic behavior are purposely being simplified as much as possible.

At time 0, it comes as a complete surprise when $\delta$ changes to $\delta'$. Suppose that after $t = 0$ everyone believes that the effort required to obtain good $i$ has jumped up discontinuously to level $\bar{e}$ ($> e'$) and is declining linearly back down to level $e'$ during the transition period $[0, T']$, thereafter to remain forever at the new steady-state level $e'$. This effort profile is depicted in Figure 1. Also shown is a typical consumer's corresponding inventory stock and consumption flow trajectories.

When the shock hits at $t = 0$, the consumer is unexpectedly caught somewhere in the previously optimal inventory cycle. For all $t < 0$, inventories had been picked up and stored in batches of size $s$, to be run down at consumption rate $d$. Suppose at $t = 0$ a consumer has on hand stock $\sigma$ ($0 \leq \sigma \leq s$). In other words, this consumer made his last purchase at time

$$(40) \qquad t_0 = \frac{s}{d}\left(\frac{\sigma}{s} - 1\right).$$

From $t = 0$ on, the consumer faces the declining effort profile depicted in Figure 1. There is then an incentive to delay the next purchase time by slowing down current consumption, because the further off a purchase can be put into the future, the lower will be the effort cost of the purchase.

In the proposed adjustment, every consumer is induced to lower consumption from rate $d$ to rate

$$(41) \qquad d' \equiv d\left(\frac{s}{s'}\right)$$

until stocks run down to zero at time

$$(42) \qquad t' \equiv \frac{\sigma}{d'} = \frac{\sigma s'}{ds}.$$

From time $t'$ on to infinity, this typical consumer repeats the new steady-state cycle: inventories of size $s'$ are picked up and stored with periodicity $T'$ [equation (39)], being run down by consuming at rate $d$.

Thus, the typical consumer with inventory $\sigma$ at $t=0$ suffers a decline in consumption by the fraction $s/s'$ over his adjustment period $[0,t']$. By time $t=T'$, all consumers are synchronized in the new steady-state.

To prove that the proposed time profiles $e(t)$, $s(t)$, and $d(t)$ depicted in Figure 1 represent a consistent dynamic transition path, two conditions must be shown to hold. First, the feasibility of these inventory and consumption patterns must be demonstrated. Then it must be proved that each consumer wants to follow this prescribed pattern.

Feasibility means that at any instant the flow of supply is equal to the amount of stock being demanded per unit time. The proposed trajectory is feasible for $t<0$ and for $t>T'$ because these are, respectively, old and new steady-state policies satisfying (18)–(20). For $t\in[0,T']$, a "spreading out" effect occurs. Buyers arriving to market now are each carrying away larger stocks than before ($s'$ instead of $s$), but these same buyers are arriving less frequently because they are running down existing stock at a rate of $d'$, which is lower than $d$. It turns out that the two effects exactly cancel each other, and in aggregate buyers are taking away the same amount per unit time as in the steady state.

In the old steady state at $t<0$, a consumer with stock $\sigma$ at time 0 would have planned on next picking up stock $s$ at time

$$\tau\equiv\frac{\sigma}{d}. \tag{43}$$

After the shock, a consumer with stock $\sigma$ at time 0 will next pick up stock $s'$ at time $t'$ defined by (42). The ratio of the frequency of buyer arrivals in the old equilibrium to the frequency of buyer arrivals in the transition interval is thus

$$\frac{t'}{\tau}=\frac{s'}{s}. \tag{44}$$

Equation (44) is exactly the ratio of the stock each buyer purchases during the transition interval to the stock each buyer purchases in the old equilibrium. Thus, the total stock demand per unit time is the same, equal to supply, and the proposed trajectory is feasible.

Next it must be shown that, given the effort profile $e(t)$ depicted in Figure 1, each consumer desires to follow the prescribed consumption and stock trajectories $d(t)$ and $s(t)$. For the steady states of $t<0$ or $t>T'$ this has already been shown by (21)–(23). To show that the effort profile $e(t)$ of Figure 1 induces the depicted pattern of $d(t)$ and $s(t)$ for $t\in[0,T']$, I must make some assumptions about consumption payments and inventory costs in the transition period. I do not think that these assumptions are critical in the sense that they are being made primarily to serve as sufficient conditions for the relatively neat patterns of Figure 1 to emerge. I believe that more general assumptions of a reasonable sort would give the same qualitative features but would be harder to analyze.

In what follows, the consumer's basic decision variable is $x$, the rate of consumption from $t=0$ until the inventory stock $\sigma$ runs out, at which time it is optimal to follow the new steady-state policy. Assume that the money cost of consuming any good is pay-as-you-go. It is proposed that consumption is at rate $x$ for $0\le t\le\sigma/x$ and at rate $d$ for $t>\sigma/x$. The utility cost of following this policy over the interval $[0,T']$ is

$$\lambda px\left(\frac{\sigma}{x}\right)+\lambda pd\left(T'-\frac{\sigma}{x}\right). \tag{45}$$

Assume further that inventory storage cost is essentially the storage-cost coefficient times the new inventory level times the interval over which it operates. Some stories could be told to support this interpretation, but it is more important to emphasize that virtually any reasonable formulation would yield qualitatively similar results. The inventory-carrying cost of the proposed policy over the interval $[0,T']$ is then

$$hs'\left(T' - \frac{\sigma}{x}\right). \tag{46}$$

For notational convenience, assume that $w$ and $v$ are normalized to unity.

With $e(t)$ representing effort as a function of time, the total net utility of the proposed policy over the interval $[0, T']$ is then

$$\left(\frac{\sigma}{x}\right)U(x) + \left(T' - \frac{\sigma}{x}\right)U(d) - e\left(\frac{\sigma}{x}\right) - \left[\lambda px\left(\frac{\sigma}{x}\right) + \lambda pd\left(T' - \frac{\sigma}{x}\right)\right] - hs'\left(T' - \frac{\sigma}{x}\right). \tag{47}$$

The consumer chooses $x$ to maximize (47), yielding the first-order condition

$$\frac{\sigma}{x}U_{\iota}(x) - \frac{\sigma}{x^2}U(x) + \frac{\sigma U(d)}{x^2} + e'\left(\frac{\sigma}{x}\right)\frac{\sigma}{x^2} - \frac{\lambda pd\sigma}{x^2} - \frac{hs'\sigma}{x^2} = 0. \tag{48}$$

The expression $e'(\sigma/x)$ stands for the time derivative of the effort function evaluated at $t = \sigma/x$.

For the consumer to be choosing to follow the prescribed path of Figure 1, condition (48) must hold for the value

$$x = d' \tag{49}$$

where $d'$ is defined by (41). Substituting (49) and (41) into (48) and rearranging yields

$$e'\left(\frac{\sigma s'}{ds}\right) = U\left(\frac{ds}{s'}\right) - U(d) - \frac{ds}{s'}U_{\iota}\left(\frac{ds}{s'}\right) + \lambda pd + hs'. \tag{50}$$

As the parameter $\sigma$ varies from 0 to $s$, condition (50) defines the slope of the required effort function $e'(t)$ on the interval $[0, T']$. Since the right-hand side of (50) does not contain $\sigma$, the slope of the effort function, hereafter denoted $e'$, must be constant throughout $[0, T']$. In other words, $e(t)$ is a straight line, as depicted in Figure 1.

From applying (32) to (50),

$$e' = 0 \quad \text{for } s' = s. \tag{51}$$

In other words, a transition from an old steady state to a new steady state that is the same as the old steady state is no transition at all.

The concave function

$$U(y) - \lambda py - \frac{hs'y}{d} \tag{52}$$

has a unique maximum, where its derivative is zero, at $y = d$ defined by the first-order condition (32). In particular, this means that, for $y = ds/s'$,

$$U(d) - \lambda pd - \frac{hs'}{d}d \geq U\left(\frac{ds}{s'}\right) - \lambda p\left(\frac{ds}{s'}\right) - \frac{hs'}{d}\left(\frac{ds}{s'}\right). \tag{53}$$

Substituting (53) into (50) and rearranging yields

$$e' \leq -\frac{ds}{s'}\left[U_{i}\left(\frac{ds}{s'}\right) - \lambda p - \frac{hs'}{d}\right]. \tag{54}$$

From concavity of expression (52), the derivative with respect to $y$ of (52) is greater than zero for $y < d$. Letting $y = ds/s'$, this observation applied to (54) yields

$$e' < 0 \quad \text{for } s' > s. \tag{55}$$

Since $e'$ is continuously differentiable in $s'$, (51) and (55) imply that, when $s'$ is sufficiently close to $s$,

$$e' > 0 \quad \text{for } s' < s. \tag{56}$$

It has been shown, then, that in undergoing a change from an old steady state with a lower degree of price distortion to a new steady state of higher price distortion, there will occur a transition phase in which effort is greater and consumption is lower than in either steady state. Conversely, during a transition to a state of less price distortion, effort will be lower, and consumption will be higher.

All of this means that increased price distortion has particularly painful immediate effects as goods are squeezed out of consumption and into inventory stocks. The situation may stabilize later of its own accord to a still unsatisfactory level, but the period immediately after an increase in monetary overhang is likely to be one of especially acute shortage symptoms.

The good news is that if price distortion can be lessened then society can go on a temporary binge as unwanted inventories are worked off into extra consumption. Reversing the shortage process—by raising prices or lowering money incomes—results in a "consumption dividend" from unneeded stocks coming out of storage into general circulation.

## V. Conclusion

This paper has presented a formal model of consumer behavior under conditions of shortage. A critical role in allocating shortage goods is played by the directly unproductive activities of search and storage. It is possible to express directly these two key variables as simple functions of the underlying degree of price distortion and, therefore, to analyze easily the relation between price distortion, search effort, and goods stockpiling. "Hoarding psychology," or what might better be called "defensive hoarding," can be thoroughly analyzed as an economic phenomenon by extensions of standard economic theory. In shortage equilibrium, everyone must expend effort to locate and hoard goods because everyone else is expending effort to locate and hoard goods.

A clear theme of the paper is that price distortion and monetary overhang can present very severe threats to the normal functioning of an economy. The essential inadequacy is in the monetary domain of prices and income, not in the real economy of production and distribution. The seeming paradox of the missing Soviet soap must be resolved not by scapegoating distributors or by making token increases in soap production, which do little to alleviate the underlying problem. Instead, prices on soap must be increased, or incomes lowered, so that consumers can move toward a better state where they are not impelled to hoard large inventories. The essential issue is to remove the incentives that lead to excessive inventory stocks blocking what should be a direct flow of goods from production to consumption.

## REFERENCES

**Buchanan, James M., Tollison, Robert D. and Tullock, Gordon,** eds., *Toward a Theory of the Rent-Seeking Society*, College Station, TX: Texas A&M Press, 1980.

**Kornai, Janos and Weibull, Jorgen W.,** "The Normal State of the Market in a Shortage Economy: A Queue Model," *Scandinavian Journal of Economics*, March 1978, *80*, 375–98.

**Osband, Kent,** "Economic Crisis in a Reforming Socialist Economy," unpublished manuscript, November 1989.

**Polterovich, Victor M.,** "Problema Izmereniya Defitsitnosti Blag" [The Problem of Measuring Welfare Under Shortage], *Ekonomika i Matematicheskie Metodi*, May 1983, *19* (5), 878–91.

**Stahl, Dale O. and Alexeev, Michael,** "The Influence of Black Markets on a Queue-Rationed Centrally Planned Economy," *Journal of Economic Theory*, April 1985, *35*, 234–50.