Cowles Foundation Paper 347
Reprinted from Quarterly Journal of Economics, 85, 1971

MATERIAL BALANCES UNDER UNCERTAINTY *

MARTIN L. WEITZMAN

I. INTRODUCTION

Critical to the success of any operational short-term central plan is the assurance that anticipated supplies will be adequate to fulfill projected demands. For many centrally planned organizations and in particular for Soviet-type economies, it is the task of the so-called method of material balances to ensure that this coordination is achieved, at least on paper. For each centrally planned commodity a balance sheet is drawn up listing sources of anticipated supply (in a closed economy mainly current production plus initial stocks) and components of projected demand (intermediate consumption plus inventory accumulation plus deliveries to users outside the system). The central planners try to work things out so that total demands and supplies are about equal. In practice projected supplies usually exceed demands, an allowance normally being made for a safety factor. This is to ensure that operating reserves will be available for use during plan execution should the need arise for unanticipated rescue operations.

Intermediate consumption of industrial materials is usually estimated on the basis of material norms (input-output coefficients). If these norms are approximately correct, or if they err on the side of safety, planned output targets will probably be fulfilled. But what if the ex post input-output coefficients turn out to be higher than the projected norms? With total supplies of certain intermediate commodities inadequate to meet industrial consumption needs, shortages will occur and the industrial supply system starts operating under strain. The immediate response is a priority rationing of

* Without blaming them for whatever defects might remain, I would like to acknowledge the useful comments of M. Keren, J. Kornai, and J. M. Montias, and the kind assistance of A. K. Klevorick. The research described in this paper was carried out under grants from the National Science Foundation and the Ford Foundation.

the item or items in question that favors its most important users. Effects of plan errors are thus cushioned not only by reserves, but also by the existence of nonpriority sectors that bear the brunt of a shortage. In the long run a supply imbalance will be corrected by increasing production relative to consumption, but in the meantime some capacity will stand idle for lack of materials.

Guided by this capsule summary, we create a highly oversimplified model of a Soviet-type industrial system. The purpose is to set up a framework for studying certain issues of optimal plan formulation and execution under uncertainty. It turns out that the stylized problem being considered can be cast in the mathematical form of a classical inventory model. This is convenient because known and powerful methods of analysis can be brought to bear on its solution.

II. The Economic Environment

The industrial system we have in mind is best thought of as a conglomerate of mining, manufacturing, power, transportation, and construction. To remove the technical difficulties of dealing with an open industrial system and to make things easy to conceptualize, we imagine that the industrial system forms a closed self-sufficient subeconomy whose final products go directly into final consumption or investment.[1]

It is assumed that each commodity can be meaningfully distinguished as being either primarily a final product or mainly an intermediate material. Unfortunately more than a few industrial items are really both, a troublesome technicality we naturally choose to ignore. We imagine that our hypothetical industrial economy is divided into two sectors — final products and intermediate materials.

Ordinarily one thinks of final products as simply those commodities that leave the system, like clothing, machines, or finished construction. For the purposes of the present paper it may be better to conceptualize the final products sector as vertically integrated a few stages back, to the extent of also including those specialized intermediate materials, like cloth, machine bearings, or cement,

1. This idea is patently false in cases of outside supply, like agricultural raw materials or foreign imports. But the myth of self-sufficiency is not a bad abstraction of the industrial system as a whole, since day-to-day workings are so typically concerned with internally produced commodities. The stochastic planning problems associated with agriculture are of quite a different nature, and for present purposes this sector is best disregarded altogether. Similar comments apply to services, distribution, and foreign trade.

whose end use is clearly and directly tied up with the production of true final products. The intermediate materials sector is best thought of as consisting mainly of basic multipurpose commodities like chemicals, metals, fuel, power, and transportation that are far back in the early stages of the production pipeline. Under this interpretation intermediate materials have the distinguishing feature of being consumed in significant proportions by the intermediate materials sector itself, as well as by the final products sector.

It goes without saying that such a fuzzy demarcation could never be used for operational purposes. But for the primitive kind of model building we have in mind, the intuitive distinction outlined above is good enough.

Each of the two main sectors is thought of as further subdivided into a number of subsectors. In turn, each subsector, headed by a ministry, is made up of enterprises (firms). An enterprise produces output according to laws of production embodied in a production function. For analytic convenience we assume that all firms throughout the economy have identical production functions. That assumption is patent nonsense as an approximation to reality. But its adoption will permit us to concentrate on the main features of the problem. With that assumption we can easily solve what would otherwise be tricky problems of plan balance, ensuring that a relatively simple optimal policy can be derived.

Production functions for each firm are postulated to be of the quasi fixed-proportions type. Let firm j possess capacity $Y_n{}^j$ during period n. If we like, we can think of capacity $Y_n{}^j$ as being created by labor and capital according to the formula

$$Y_n{}^j = F_n(L_n{}^j,\ K_n{}^j),$$

with $F_n(\quad)$ a constant returns to scale capacity function common to all enterprises for period n. If intermediate materials $M_n{}^j$ are available to enterprise j, output $Q_n{}^j$ is given by

$$(1) \qquad Q_n{}^j = \min\left\{ Y_n{}^j, \frac{M_n{}^j}{\theta_n} \right\},$$

where θ_n is the common input-output coefficient of each enterprise for period n.[2]

2. If intermediate materials $M_n{}^j$ were disaggregated into more specific individual types, a fixed-proportions production relation analogous to (1) but containing as arguments the various specific materials would probably be too rigid a specification. With certain commodities in tight supply, users would be encouraged to keep up their own outputs by using less scarce substitute materials. But if materials as a whole are in short supply, there is relatively little room for maneuverability. Capital and labor may be fair substitutes for one another in the creation of capacity, but in most production

III. Mechanics of Planning

Plans are formulated instantaneously at times 0, 1, 2, ... A plan is executed during the unit of time directly following formulation, known as the plan period. We follow the convention that a plan formulated at *time* $n-1$ is executed during *period* n.

We suppose that the industrial economy has the potential of producing gross output Y_n during plan period n. This capacity estimate might be derived as $F_n(L_n, K_n)$, where L_n is the total man hours, K_n is the aggregate capital stock, and $F_n(\ \ )$ is here interpreted as an aggregate capacity function.[3]

For the purposes of this paper operational planning is seen as having an essential putty-clay aspect. The putty-clay view of planning is based primarily on the notion that labor and capital are far more easily shiftable at the time of plan formulation than during the immediately following period of execution.

At time $n-1$ or before, total capacity Y_n can be split up in any proportions between intermediate materials capacity I_n^* and final products capacity F_n^*, so long as $I_n^*+F_n^* \leqq Y_n$. During plan execution, the planners are stuck with the chosen proportions in the sense that capacities I_n^* and F_n^* represent *maximum* allowable gross outputs of intermediate materials and final products during period n. These maximum levels would not both be simultaneously operative if relative to demand there were an insufficient supply of intermediate materials, due perhaps to poor planning or unforeseen difficulties.

The scenario that has been presented must be visualized as an abstraction at best. In the real world, plans are not formulated instantaneously at a moment of time just before plan execution. Planning takes time. Sometimes the plans for a period are incomplete when even a significant part of the period has elapsed; at other times, except for some quick patch-up work, plans for a given period have essentially been formulated several periods before.[4]

Even more disconcerting is the simultaneous existence in most centrally planned economies of several overlapping operational plans of varying length (chiefly the annual, quarterly, and monthly

processes either one is a poor substitute for materials. This fact is reflected in the assumption of a zero elasticity of substitution between capacity and aggregate materials in (1).

3. $F_n(\ \ )$ will be an aggregate capacity function provided that capital and labor have been efficiently combined in the same proportions for each enterprise.

4. Both of these are frequently the case with quarterly or monthly plans. Only in preparing the annual *tekhfinprom* plan is a careful job of material balances done, although the method is applied on paper to quarterly and monthly plans as well.

plans). Which plan do we have in mind as a prototype for our model economy? The answer to this question is bound up with our view of the relevant putty-clay horizon. The appropriate period should be long enough to make sense of the notion that (within a relevant range) capacities can be adjusted beforehand. Capital need not be literally freely shiftable, but there should be enough flexibility in some underlying factors (perhaps labor and newly installed capital) to permit such capacity changes of a reasonable size as might be required by the plan. On the other hand, the plan period must be sufficiently short to justify the imposition of capacity constraints *during* plan execution. About the only sure thing is that, whatever period we choose, the present model will exaggerate both preplan flexibility and intraplan rigidity.[5]

Even the real world calculation of overall economic capacity would have to be crude. Y_n is just an imperfect reflection of underlying shared resources fixed in the short run, and of decisions about how hard to work labor and capital. Unfortunately, the planners' notion of capacity is partially determined as the outcome of a bargaining process between the center and the periphery; not surprisingly, we find it convenient to disregard this aspect altogether.

IV. Material Reserves

In plan formulation and execution, a significant role is invariably played by the stock of intermediate materials held by central agencies for dealing with unforeseen contingencies. The Soviets often use the word "reserves" in a very broad sense, denoting virtually any potential for increasing output, including such intangibles as efficiency or even inspired improvisation. As we will use the term, material reserves (or just plain reserves) will mean physical stock of warehoused commodities held and distributed by or for the supply departments of Gosplan and the ministries as part of the industrial materials supply system.

In Soviet parlance material reserves additionally include national defense and natural disaster state stockpiles, which we exclude from consideration. Also excluded from our usage of the term "ma-

5. If forced to choose, I would personally pick the quarterly plan as the best single compromise. However, the monthly plan is really more relevant for some industrial materials over which close central control is maintained, while the yearly plan is more appropriate for other, more loosely controlled commodities. A complicating factor is that fulfillment of the annual and quarterly plan tends to be a more significant success indicator for the ministry, whereas the quarterly and monthly plan targets are frequently more important to the enterprise.

terial reserves" are the ordinary day-to-day production inventories held by the enterprises. This class of raw materials, semifabricates, and finished products is conceptualized as being so closely tied up with normal production processes that removal would impair production immediately or within a short time. Enterprises are forbidden to hold in their own name material reserves as we are using the term, and are restricted by law to short-term production inventories. Managerial hoarding of reserves certainly occurs, as an ample number of anecdotes bear witness, but for our purposes material reserves are best considered held by or for agencies higher than the enterprises.

The basic idea behind our usage of the term "material reserves" should be clear even though in specific cases it might be difficult to judge the degree of overlap with state stockpiles at one end or production inventories at the other. As is made abundantly clear by the Soviets themselves, the purpose of holding material reserves is to be able to remedy such branch "disproportions" as may arise during plan execution.

V. Plan Formulation and Material Balances

In the context of the present model, planning material balances is an especially simple procedure.[6] Suppose that for period n the true input-output coefficient θ_n is not yet known exactly, but that μ_n, an approximate norm of materials consumed per unit of output, is used instead for plan construction. The proxy input-output coefficient μ_n might be estimated as an average input per unit of output over time. Let R_{n-1} be the stock of material reserves on hand at time $n-1$.[7] We take I_n^* and F_n^* as the period n target levels of intermediate materials and final products, respectively.

Total supplies of materials for the coming period are $I_n^*+R_{n-1}$. Anticipated intermediate consumption is $\mu_n(I_n^*+F_n^*)$. The planners try to make sure that

$$I_n^*+R_{n-1}>\mu_n(I_n^*+F_n^*).$$

If positive, the difference

$$I_n^*+R_{n-1}-\mu_n(I_n^*+F_n^*)$$

6. We assume that all commodities are centrally planned. Technically this would only be true of the so-called "funded" and "planned" commodities. However, all of the important industrial materials are covered by these two categories.

7. It is impossible to hold material reserves of electricity or transportation as such. But close substitutes are available in the form of fuels and emergency standby equipment, the latter usually existing in a semiretired state.

is a safety factor that is often built in to ensure that sufficient material reserves will be available throughout the plan execution phase. Capacity constraints require $I_n^* + F_n^* \leq Y_n$. Real world plan formulation is of course much more difficult, in part because intermediate consumption requirements are *not* independent of output composition.

After it is composed the plan is handed out to the ministries, and through them is broken down into enterprise plans and dispersed to the firms. For convenience we choose to ignore the subsequent bargaining over assignments, which is an important part of actual planning.

Real plan targets are set on many aspects of performance, including productivity or employment of labor and materials, profits, and costs. But there is no doubt that the most important target on both the ministerial and enterprise level is the output quota.[8]

VI. Effects of Uncertainty

Uncertainty can interfere with plan fulfillment in many ways. We divide the obstacles to accurate planning by the method of material balances into two broad categories. The first, on which this paper concentrates, is Gosplan's inaccurate knowledge of input-output coefficients. There is no doubt that this is a significant and recurring problem in Soviet-type planning. It is caused mainly by an inability to represent accurately complicated and continually changing production processes with a system of simple aggregate norms. There are also unforeseen differences between various enterprises and ministries arising from the uneven quality of equipment, input materials, and labor skills, from incentives to distort, and from the unknown distribution of hoarded materials. In addition, errors are frequently compounded by faulty aggregation.

A second cause of disturbances, conceptually somewhat different in nature, is the unforeseen contingencies that arise during plan

8. Let all ministries and enterprises be called *micro-units*. Micro-unit output quotas are for end products of the unit only and would be strictly net of those end products simultaneously consumed as inputs by the producing unit. Gosplan is interested in the amount of sulfuric acid made available to programmed users outside the chemical sector, not in the amount internally consumed by the chemical ministry in the manufacture of other industrial acids. But note what happens when micro-units having input-output relations with each other are aggregated together into an artificial intermediate materials sector (which produces a single product functioning as both input and output). Output targets on the micro level framed *net* of self-produced inputs become blown up into an intermediate materials output target that is *gross* of self-produced inputs.

execution. In this category are such diverse events as unexpected equipment failure, sudden change in final demand due to political or other events, unanticipated speed-ups or lags in delivering expected capacity, or general supply foul-ups due to, e.g., adverse weather.

Actually, most disturbances of the second type can be translated into the language of norm uncertainty. The overriding importance of the output quota to a certain extent normalizes enterprise output levels. With enterprise bonuses and prestige linked primarily to output quota fulfillment, managers are in effect encouraged to use the enterprise fund or to cut into current profits in order to substitute labor for capital (with overtime, increased shifts, and "rush work" if necessary), or to take other emergency measures to ensure holding up their end by fulfilling the output target. This built-in flexibility in being able to adjust partially the timing and quantity of factors semipermanently attached to the enterprise tends to stabilize output somewhat at the expense of making input requirements more variable. Such adjustments as occur will usually be automatic, without any explicit orders from higher up, and will tend to be therefore more or less instantaneous.

Of course, the authorities will frown on using up too much labor or materials, or not making a high enough profit, and the bonuses will be correspondingly lower. But within limitations, committing these offenses will generally be preferable to cutting output below the target level. Should extra inputs over the budgeted amounts be required to meet output quotas, there will usually be no problem for an important enterprise or one supplying deficit materials to obtain supplementary *nariady* (procurement orders). At least a lower priority "nonplan allocation order" will usually be issued to any organization with a good story. Occasionally needed materials can be illegally obtained on the basis of pure *blat* (pull) alone. Since directors are almost always willing to sacrifice cost and profit targets to meet the output quota, the critical operational question is whether extra supplies *exist* anywhere in the system. If they do, the *tolkachi* (pushers) from some organization will uncover them.

On the other hand, even an abundance of intermediate materials will not ordinarily lure enterprise managers or industrial ministers into overfulfilling output quotas by a conspicuous margin. Due to the operation of an almost universal ratchet principle of planning, the formation of next period's plan targets will start off with this period's performance as a point of departure. Benefits, material or otherwise, increase tremendously for plan fulfillment but not very much more per degree of overfulfillment. An overzealous performance in any given period can have disastrous long-run effects on the organization in question.

We take as our point of departure the approximation that output targets will be exactly fulfilled if the necessary input materials are available. One would think that, as inventories are run down or pile up, the center would revise upward the production targets of, respectively, intermediate materials or final products. We allow such changes to become effective starting *next* period. The whole idea of the model is that plan periods are sufficiently short to make it difficult to increase outputs after the plan has been formulated. This comes about primarily because it is difficult to shift underlying resources on short notice. The tendency toward short-run nonshiftability is reinforced because the time that it takes for the center to notice imbalances, draw up new plans, and have target revisions reach down to the level of the enterprise, introduces a lag of its own. As previously indicated, the relevant abstraction is that plan targets I_n^* and F_n^* are upper limits on gross outputs of intermediate materials and final products during period n.

VII. Plan Execution and the Costs of Incorrect Planning

Let θ_n, with $0<\theta_n<1$, be the true value of the economy-wide input-output coefficient in period n. During plan formation at time $n-1$, only the distribution of the random variable θ_n is known.[9] Let I_n and F_n represent actual period n gross outputs of, respectively, intermediate materials and final products. If

$$I_n^*+R_{n-1}\geqq\theta_n(I_n^*+F_n^*), \tag{2}$$

all output targets will be fulfilled and

$$\begin{aligned} I_n&=I_n^*,\\ F_n&=F_n^*. \end{aligned} \tag{3}$$

Stocks of strategic reserves will change from R_{n-1} at time $n-1$ to

$$R_n=R_{n-1}+I_n^*-\theta_n(I_n^*+F_n^*) \tag{4}$$

at time n.

If, on the other hand,

$$I_n^*+R_{n-1}<\theta_n(I_n^*+F_n^*), \tag{5}$$

9. Only analytic convenience impels us to accept the idea that the input needs of both sectors are described by the *same* random variable. Two different, more or less independent random variables would yield a much more realistic description. But employing more than one random variable clutters up the analysis and destroys the simplicity of an optimal policy without really changing the basic ideas of the model.

an ex post plan inconsistency exists. Not all planned output quotas can be simultaneously fulfilled because someone has to go short of inputs. Smooth operations of the materials supply system will start to break down as reserves dwindle and rationing becomes necessary.

The ex post maximum attainable final product, F_n, is the solution to the linear programming problem:

$$\begin{aligned} &\max F \\ \text{subject to } &I \leqq I_n^* \\ &F \leqq F_n^* \\ I+R_{n-1} &\geqq \theta_n(I+F) \\ &I,\ F \geqq 0. \end{aligned}$$

The production function implicit in the above formulation is based on (1).

With (5) holding and $0<\theta_n<1$, the solution of this linear programming problem is

$$I_n = I_n^*$$

$$(6)\qquad F_n = \frac{I_n^* + R_{n-1}}{\theta_n} - I_n^*.$$

If shortages occur in intermediate materials it will not pay to cut inputs from sectors producing the deficit materials. Such action would only magnify the deficit via a multiplier effect.[1]

The linear programming problem need never be solved formally to obtain the solution I_n, F_n. The optimal solution will automatically be generated and enforced if an obvious rationing procedure is followed. As reserves decrease, pressure will mount for producers of input materials in short supply to keep up their target outputs (and even to increase outputs, a possibility we disallow until next period). Producers of deficit materials will tend to become priority users of rationed inputs, since they are holding up the system. Producers of final products will get whatever materials remain. Note the critical, if simple, role of reserve levels in determining who is a priority user. At the end of the plan period no material reserves will be left,

$$(7)\qquad R_n = 0,$$

since all intermediate goods will have been used up in bailing out sagging enterprises.

1. This conclusion, so obvious in the present framework, has an interesting operational generalization to the issue of optimal rationing in a multisector model. Cf. M. Manove, "A Theory of Administrative Planning in Soviet Type Economies," Ph.D. thesis, Massachusetts Institute of Technology, 1969.

In our formulation, the cost of overtaut planning is the loss of capacity $F_n^* - F_n$ and the reduction of final product that it entails. Too little provision for intermediate goods can cause a plan to break down because the economy is temporarily frozen into a situation of insufficient capacity for intermediate materials production. On the other hand, the cost of overloose planning is the final product lost by the failure to convert abundant material reserves into final products. In both extreme cases losses arise because capacity is temporarily locked into inappropriate plan proportions and cannot be instantaneously shifted to assist the overburdened sector. Optimal plans are a balance between the two extremes.

VIII. Short- and Long-Term Plans

We have been speaking of short-term plans as if each one were independent of the others. In fact, current operational plans are loosely embedded in the long-term plan. The latter is a nonoperational plan that outlines the general dimensions of economic growth over a period of several years. We assume that as of now (time zero) the long-term plan embodies current growth strategy, although in reality the plan document may or may not formally be brought up to date at any given time.

Let N short-term planning periods remain until the expiration of the current long-term plan. To a first approximation we take capacities $\{Y_n\}^N_{n=1}$ as exogenously determined. Economic growth ensures that $Y_n \leqq Y_{n+1}$ for all n. The appropriate planning discount or interest rate r is also treated as given.[2] Since it is unlikely to be known with any degree of accuracy, assuming r to be constant during the long-term plan is worth the convenience it creates. The short-term planners are postulated to operate by treating $\{Y_n\}$ and r as data outside their ability to control.[3]

2. Until recently, in Soviet economic theory and practice, shadow interest rates have been explicitly employed only rarely. But very close substitutes have long been utilized in many of the standard investment criteria. Obviously any rationally planned economic society must have and make use of some notion of the trade-off between present and future income. We merely take r as a crude proxy for the rate of return to social savings.

3. The reasoning behind the exogenous status of $\{Y_n\}$ and r can be explained by considering overall capacity as an aggregate function of total capital and labor, $Y_n = F_n(L_n, K_n)$. The size of the labor force is more or less exogenously determined at any time. The quantity of available man hours additionally reflects decisions about how hard people should work, also treated exogenously. Capital stock is cumulated out of investments going back for many years. In that length of time the law of large numbers will tend to smooth out possible effects of plan performance fluctuations. These kinds of disturbances taken as a whole are rarely so violent as to alter in-

Let $\alpha \equiv \frac{1}{1+r}$. The discount factor α can be identified with the marginal social rate of substitution and also with the marginal rate of physical transformation between goods produced in any two periods of the current long-term plan. The factor α^{n-1} will be used to convert the costs and benefits incurred during period n to a common base at time zero.

It is assumed that the random variables $\{\theta_n\}$ are independently identically distributed with probability density function $f(\theta)$.[4] On a priori grounds we restrict $f(\theta)$ to take on positive values only for θ between c and C, where $c>0$ is the minimum conceivable and $C<1$ is the maximum conceivable input-output coefficient. Thus,

$$\int_{c}^{C} f(\theta)\,d\theta=1. \tag{8}$$

IX. Optimal Planning and Dynamic Programming

The overall objective in formulating short-term plans knit together into a long-term plan is taken to be a maximand that is a sum of two parts. The first is the expected discounted value of final products produced during the current long-term plan,

$$\sum_{n=1}^{N} \alpha^{n-1}F_n.$$

The second part is the expected discounted value of the strategic reserves that will be left over to be used during the next long-term plan. The value of the inventory bequest R_N will be treated simply as $\alpha^N R_N$, consistent with our previous interpretation of α as the appropriate discount factor.

Plan targets for period n are set at time $n-1$. At that time the stock of strategic reserves is R_{n-1}, treated as an inheritance from the past about which nothing can be done.

Let $\Psi_n(R)$ represent the discounted expected value of an opti-

vestments in a given year anyway. This is especially true in Soviet practice because productive investment is insulated from input shortages at the expense of other sectors. The same remarks can be used to justify the exogenous treatment of the interest rate, which is roughly taken as representing the marginal product of capital. The assumption that interest rates are constant can be lifted at the expense of destroying part of the simplicity of an optimal policy.

4. In fact, ex post intermediate material norms are undoubtedly not distributed independently or identically from one time to another for a wide variety of reasons. Nevertheless, it is felt that the convenient assumption to the contrary is not a bad first approximation in the present context. Arbitrarily distributed material norms could be handled only at a cost of generating more complicated optimal policies.

mal policy starting at time $n-1$ (period n) with material reserves R. At time N (period $N+1$),

$$\text{(9)} \qquad \Psi_{N+1}(R) \equiv a^N R.$$

Define $Z \equiv \frac{I+R}{I+F}$ to be the ratio of total planned supply of intermediate materials available as inputs during the plan period divided by the total planned output of the economy. If the true input-output coefficient θ turns out to be less than Z, all planned output targets will be fulfilled and (3) and (4) hold. With θ greater than Z, the plan breaks down and improvised allocations (6) and (7) take place. For periods $n=N, N-1, \ldots, 1$, $\Psi_n(R)$ is therefore recursively defined by the dynamic programming equation

$$\text{(10)} \qquad \Psi_n(R) = \max_{\substack{I+F \leqq Y_n \\ I,\ F \geqq 0}} \left\{ \int_0^Z a^{n-1} F \, f(\theta)\, d\theta + \int_Z^1 a^{n-1} \left(\frac{I+R}{\theta} - I \right) f(\theta)\, d\theta + \int_0^Z \Psi_{n+1}(I+R-\theta(I+F)) \, f(\theta)\, d\theta + \int_Z^1 \Psi_{n+1}(0) \, f(\theta)\, d\theta \right\},$$

where $Z \equiv \frac{I+R}{I+F}$. The first two terms inside the maximand of (10) represent for a given value of Z the expected discounted final output in period n (from (3) and (6)). The last two terms represent the discounted expected value of an optimal policy from periods $n+1$ to N taking into account (from (4) and (7)) the reserves left over at the end of period n.

An optimal policy at time $n-1$, given reserves R_{n-1}, consists of the targets $I_n^*(R_{n-1})$, $F_n^*(R_{n-1})$, which maximize the right-hand side of (10) subject to the constraints and for $R = R_{n-1}$.

The following theorem characterizes the form of an optimal policy for all n:

An optimal policy can be described by a single critical number [5] *s defined as the unique solution of the equation*

5. Note that an optimal policy that depends on a *single* critical number has been obtained only at the expense of making several simplifying assumptions. Under more general conditions we would not expect and could not obtain such a sharp characterization of an optimal policy.

$$(11) \qquad -1+\int_s^1 \frac{1}{\theta} f(\theta)\,d\theta+\alpha \int_0^s f(\theta)\,d\theta=0.$$

If $R_{n-1}<sY_n$, $I_n^*=sY_n-R_{n-1}$. *If* $R_{n-1}\geqq sY_n$, $I_n^*=0$.
In either case, $F_n^*=Y_n-I_n^*$.

As will be indicated in the Appendix, the only situation likely to be met in practice is $R_{n-1}<sY_n$ with the corresponding rule $I_n^*=sY_n-R_{n-1}$, $F_n^*=Y_n-I_n^*$. This rule has an obvious interpretation in terms of a stock adjustment principle. Let μ represent the planners' notion of an "average" value of θ that is used as a norm for projecting input needs.

The proposed planning rule is equivalent to

$$(12) \qquad I_n^*=\mu Y_n+[\lambda Y_n-R_{n-1}],$$

with $\lambda\equiv s-\mu$. Projected consumption of intermediate materials is μY_n. We interpret λ as the desired inventory norm; λY_n is then the preferred safety level of reserves during period n. The quantity $[\lambda Y_n-R_{n-1}]$ is the difference between current desired and current actual reserves. Production of intermediates is targeted to cover anticipated materials consumption and to bring reserves to the desired safety level. To take the analogy one step further, μY_n consists of those already committed input materials for which *nariady* have been issued to the enterprises, whereas $[\lambda Y_n-R_{n-1}]$ is the current uncommitted output of intermediate materials set aside for possible emergency uses by plan administrators.

Equation (11), which determines the optimal value of s, the planned ratio of available inputs to outputs, has an interesting economic interpretation. In any period consider transferring a small unit of capacity from F^* to I^*, provided both F^* and I^* are positive. If $\theta\leqslant s$, a unit of final output will simply be lost so that the expected loss of final product is just the probability that θ is less than s, $\int_0^s f(\theta)\,d\theta$. If $\theta>s$, we are in a regime where (6) holds; the indirect gain in final product of $1/\theta$ offsets the direct loss of a unit of final output so that the expected net gain of current final product would be $\int_s^1 \left(\frac{1-\theta}{\theta}\right) f(\theta)\,d\theta$. The expected increase in next period's stock of inherited reserves is just the probability that θ is less than s, whose expected value discounted back to this period is $\alpha\int_0^s f(\theta)\,d\theta$. If s is to be optimal and all capacity is being utilized, potential gains should exactly offset potential losses, resulting in equation (11).

Because the proof of the theorem is somewhat technical, it is reserved for the Appendix.

X. Sensitivity Analysis

The remainder of the formal part of this paper is a comparative statics analysis of the effects of various parameter changes on optimal plan target levels.

The easiest effect to analyze is that caused by changes in the discount rate. Differentiating (11) with respect to a and collecting terms, we find that

$$\frac{ds}{da}=\frac{\int_0^s f(\theta)\,d\theta}{f(s)\left(\frac{1}{s}-a\right)}>0.$$

The positive sign of $\frac{ds}{da}$ is intuitively obvious. As leftover reserve stocks are valued higher because of lower shadow interest rates, an incentive is created to increase the production of intermediates relative to final products. Other things being equal, lower interest rates should be associated via higher s with a decreased likelihood that the industrial supply system breaks down and goes over to the critical phase where deficit materials are rationed.

This observation may have some bearing on the issue of "optimal tautness" in planning. Plan "tautness" can be described in very broad terms as the degree of "supply tension" under which economic units operate. In the present model some more or less equivalent measures of the degree of tautness in a formulated plan might be the target level of reserves as a fraction of capacity, $\lambda\equiv s-\mu$, or the probability of precipitating a supply crisis, $\int_s^1 f(\theta)\,d\theta$, or the percent of capacity expected to stand idle for lack of materials, $\int_s^1\left(1-\frac{s}{\theta}\right)f(\theta)\,d\theta$.

It has been observed that, other things being equal, plans are frequently more taut in the earlier than in the later stages of industrialization.[6] While there are undoubtedly a wide variety of reasons for this phenomenon, we merely record the existence of an economic rationale stemming from the present model. In the early stages of development when the marginal product of capital is rela-

6. H. Hunter concentrates primarily on the hortatory effects of high targets ("Optimal Tautness in Development Planning," *Economic Development and Cultural Change*, July 1961, pp. 561–72). The analysis presented here is not meant to suggest that motivational (or other) aspects of target setting are not important, especially in the early stages of development.

tively high, taut planning makes good economic sense. With a high implicit interest rate it is more important to use up available stocks of materials *now* at the expense of next period's initial stocks. This enlarges the chance that intermediate materials will become rationed deficit commodities and that some capacity will stand idle. Within the present framework the optimal tautness of plan targets should be eased over time as the interest rate declines.

In analyzing the effects of uncertainty on the target level of strategic reserves, we start with the simplest case. Were θ known in advance with perfect certainty to be equal to μ, s would also be set at μ. In a world of perfect certainty, $\lambda=0$ — there is no need for reserves.

Whether the introduction of uncertainty around μ will make s greater or less than μ would depend, among other things, on the values of θ allowed by its probability distribution. With θ relatively small, it would probably pay to keep a positive safety reserve because large capacity losses would accompany input shortages (as we will show in the next two paragraphs, this is the case of practical interest). On the other hand, if θ tends to be high and the discount rate α is low, the cost of breakdowns diminishes and the so-called "safety reserve" might even be negative. These ambiguous conclusions come from examining the equation

$$(13) \qquad \int_c^s (1-\alpha)\ f(\theta)\,d\theta = \int_s^C \left(\frac{1}{\theta}-1\right) f(\theta)\,d\theta,$$

which is equivalent to (11). The term $(1/\theta-1)$ could be greater or less than $(1-\alpha)$ depending on the values of θ and α, and this determines the magnitude of s relative to μ.

Now let μ be taken as the median of the probability density function $f(\theta)$. The value of s in (13) will be compared with the perfect certainty case $s=\mu$. Suppose that

$$(14) \qquad C \leqq \frac{1}{2-\alpha},$$

with C the maximum conceivable value of θ. Condition (14) should certainly hold for any realistic values of C and α (remember that the short-run plan period we envision is probably less than a year in duration, making α higher than an annual discount factor).

With (14) holding, we can see from (13) that $s>\mu$. The presence of uncertainty has a *positive* effect on the size of the material reserves target ratio $(s-\mu)$. This accords with the commonsense interpretation of material reserves as insurance against risk.[7]

7. Unfortunately, definite results can be obtained neither for general

XI. Concluding Remarks

A large number of the assumptions behind the present model have been motivated primarily by analytic convenience as opposed to realism. In many cases their replacement by alternatives closer to reality would result in a more intricate model with more complicated optimal policies. But such a model would still display many of the same basic features as the present one.

Unfortunately, this is only partially true of our two-sector approximation of the planned economy. Planning in a multisector world has a flavor all its own that is difficult to capture with a mere two sectors. For one thing, causality becomes very complicated with many goods — everything depends on everything else and it is not so easy to slap together consistent plans. For another thing, shortages or surpluses will usually exist in certain specific inputs and not for intermediate materials in general; an important feature of emergency resource allocation during plan breakdowns concerns the possibility of short-run substitution between similar materials. All of these kinds of notions are lost in a two-sector model.

On the other hand, the two basic conclusions of the present model, listed below, would surely generalize to more complicated situations. And these two propositions are undoubtedly easier to visualize in the present more simplified framework.

1. Short-term plan proportions should be balanced between the production of intermediate and final goods. Resources are wasted with either overtaut or too loose plan proportions.

2. The level of material reserves plays an important role as a signaling device. In the short-run plan execution phase, when capacity is more or less fixed, inventory deficits correctly indicate the priority of inputs going to producers of deficit commodities. When capacities can be altered in the longer run, current material reserve levels point out the optimal direction of next period's plan proportions.

Appendix: Proof of the Form of an Optimal Policy

To prove the theorem, we apply with appropriate modifications the kind of standard dynamic programming arguments used in de-

distribution changes affecting central tendency (e.g., increases or decreases in stochastic dominance) nor for general distribution changes altering the degree of dispersion (e.g., mean preserving spreads and contractions). Some numerical experiments with a uniform distribution in (11) indicate surprisingly high values of s close to the upper limit C and much higher than the mean or median. But this result may be at least partially due to the special features of the uniform probability density function.

termining the form of an optimal policy for inventory models. Indeed, the present model is mathematically equivalent to the familiar dynamic inventory model with linear ordering costs and a specific form of nonstationarity.[8]

Proof:

Before going on to the main part of the proof, we give a preliminary argument to show that $I_n^* + F_n^* = Y_n$. This part of the theorem is intuitively obvious, since otherwise capacity is being needlessly wasted. We will then be able to replace the inequality $I + F \leqq Y_n$ by the equality $I + F = Y_n$ without loss of generality, because the inequality will always hold as a full equality in any optimal policy. Reducing a nominally two-variable problem to one involving a single unknown represents a considerable notational simplification that it behooves us to employ as early as possible.

Working with (10), one can easily verify that if $\Psi_{n+1}(R)$ is monotone increasing in R, so is $\Psi_n(R)$. Obviously $\Psi_{N+1}(R)$ is monotone increasing in R. Hence, so too is $\Psi_n(R)$ for all n.

Suppose that (10) holds with $I_n^* + F_n^* < Y_n$. Partially differentiating the right-hand side of (10) with respect to I, one obtains after simplification

$$\int_Z^1 \left(\frac{1}{\theta} - 1 \right) f(\theta)\,d\theta + \int_0^Z \Psi'_{n+1}(I + R - \theta(I + F)) \quad (1 - \theta) \quad f(\theta)\,d\theta,$$

which is obviously positive.[9] Thus, I_n^* and F_n^* such that $I_n^* + F_n^* < Y_n$ cannot possibly maximize the right-hand side of (10). Henceforth we work with $I + F = Y_n$.

Define M as follows:

$$M \equiv R + I.$$

M represents the total materials available for intermediate consumption. For notational convenience,

$$m \equiv \frac{M}{Y_n}.$$

Let $L_n(M)$ be defined by

$$L_n(M) \equiv Y_n \int_0^m f(\theta)\,d\theta + \int_m^1 \frac{M}{\theta} f(\theta)\,d\theta. \tag{15}$$

Finally, $G_n(M)$ is determined by the equation

$$G_n(M) \equiv \alpha^{n-1}[-M + L_n(M)] + \int_0^m \Psi_{n-1}(M - \theta Y_n) \quad f(\theta)\,d\theta + \int_m^1 \Psi_{n+1}(0) f(\theta)\,d\theta. \tag{16}$$

8. There is a prolific literature on this model. See, e.g., R. Bellman, I. Glicksberg, and O. Gross, "On the Optimal Inventory Equation," *Management Science*, Vol. 2 (1955), pp. 83–104; K. J. Arrow, S. Karlin, and H. Scarf, *Studies in the Mathematical Theory of Inventory and Production* (Stanford University Press, 1958); or S. Karlin, "Dynamic Inventory Policy with Varying Stochastic Demands," *Management Science*, April 1960, pp. 731–58.

9. There is a little hand-waving going on here because we don't yet know that $\Psi_{n+1}(R)$ has a derivative, but it is trivial to give a rigorous proof by induction.

Equation (10) is equivalent to

$$(17) \qquad \Psi_n(R) = \max_{R \leqq M \leqq R+Y_n} \{a^{n-1}R + G_n(M)\}.$$

The value of M maximizing the right-hand side of (17) subject to the stated constraint is denoted $M_n^*(R)$.

The proof is by induction on n, working backwards from $n=N$ to $n=1$. For each n, propositions (i) and (ii) will be proved:

(i) There is a positive critical number S_n such that an optimal policy requires $M_n^*=S_n$ for $R_{n-1}<S_n$ and $M_n^*=R_{n-1}$ for $R_{n-1} \geqq S_n$. Furthermore, $S_n = sY_n$ where s is defined as the unique positive solution of (11).

(ii) The first derivative $\Psi_n'(R)$ exists and is continuous for all nonnegative R, and $0<\Psi_n'(R) \leqq a^{n-1}$. Furthermore, $\Psi_n'(R)=a^{n-1}$ for $0 \leqslant R \leqslant S_n$. The second derivative $\Psi_n''(R)$ exists and is nonpositive for all R except possibly the point $R=S_n$; however, left- and right-hand bounded derivatives exist at that point.

We start with $n=N$. Using (9), we find that equations (16) and (17) become

$$(18) \qquad G_N(M) = a^{N-1}[-M+L_N(M)] + a^N \int_0^m (M-\theta Y_N)\, f(\theta)\, d\theta,$$

$$(19) \qquad \Psi_N(R) = \max_{R \leqq M \leqq R+Y_n} \{a^{N-1}R + G_N(M)\}.$$

Let S_N be a solution of

$$G_N(S_N) = \max_{M \geqq 0} G_N(M).$$

Differentiating $G_N(M)$ yields

$$G_N'(M) = a^{N-1}\left[-1+\int_m^1 \frac{1}{\theta} f(\theta)\,d\theta + a\int_0^m f(\theta)\,d\theta\right],$$

$$(20) \qquad G_N''(M) = \frac{a^{N-1} f(m)\left[-\dfrac{Y_N}{M}+a\right]}{Y_N}.$$

From (20), $G_N''(M) \leqq 0$ for all M. Since $\lim_{M \to 0^+} G'_N(M) > 0$ and $\lim_{M \to \infty} G_N'(M) < 0$, we conclude that $0 < S_N < \infty$, and that S_N satisfies

$$G_N'(S_N) = 0,$$

which can be rewritten as

$$(11) \qquad -1+\int_s^1 \frac{1}{\theta} f(\theta)\,d\theta + a\int_0^s f(\theta)\,d\theta = 0$$

for $s \equiv \dfrac{S_N}{Y_N}$.

Only a value of s between c and C will satisfy (11). Note that s is unique because differentiating (11) with respect to s yields

$$(21)\qquad f(s)\left[-\frac{1}{s}+\alpha\right],$$

which is negative for $c<s<C$ and zero for other values of s.

In view of the unimodality of $G_N(M)$, it is apparent that the policy described by (i) solves (19) except possibly for the inequality $M\leqq R+Y_N$. However, this inequality will be satisfied automatically by $M_N^* = \max\ \{R,\ S_N\}$ since $S_N=sY_N<Y_N$.

From (i) and (19) we have

$$\Psi_N(R)=\begin{cases}\alpha^{N-1}R+G_N(S_N) & R<S_N\\ \alpha^{N-1}R+G_N(R) & R\geqslant S_N\end{cases}.$$

Differentiation yields

$$(22)\qquad \Psi_N'(R)=\begin{cases}\alpha^{N-1} & R<S_N\\ \alpha^{N-1}+G_N'(R) & R\geqslant S_N\end{cases},$$

where $-\alpha^{N-1}<G_N'(R)\leqslant 0$ for all $R\geqslant S_N$. $\Psi_N'(R)$ is continuous at $R=S_N$ because $G_N'(S_N)=0$. The second derivative $\Psi_N''(R)$ exists everywhere except possibly at the point $R=S_N$, where, however, left- and right-hand bounded derivatives exist.

From (20) and (22),

$$\Psi_N''(R)\leqq 0$$

for all R except possibly at $R=S_N$.

Assuming now that (i), (ii) have been proved for period $n+1$, we show that they hold for period n. Equations (16) and (17) define $G_n(M)$ and $\Psi_n(R)$ · S_n is defined as a solution to

$$G_n(S_n)=\max_{M\geqq 0} G_n(M).$$

Differentiating $G_n(M)$ yields

$$G_n'(M)=\alpha^{n-1}\left[-1+\int_m^1\frac{1}{\theta}\,f(\theta)\,d\theta\right]+\int_0^m\Psi'_{n+1}(M-\theta Y_n)\ f(\theta)\,d\theta$$

$$G_n''(M)=\frac{-\alpha^{n-1}\ f(m)}{M}+\frac{\Psi'_{n+1}(0)\ f(m)}{Y_n}+\int_0^m\Psi''_{n+1}(M-\theta Y_n)\ f(\theta)\,d\theta$$

$$=\frac{\alpha^{n-1}f(m)\left(-\dfrac{Y_n}{M}+\alpha\right)}{Y_n}+\int_0^m\Psi''_{n+1}(M-\theta Y_n)\ f(\theta)\,d\theta.$$

Thus $\lim_{M\to 0^+} G_n'(M)>0$, $\lim_{M\to\infty} G_n'(M)<0$, and $G_n''(M)\leqslant 0$ for all $M\geqslant 0$. It follows that $0<S_n<\infty$, and

$$G_n'(S_n)=0.$$

Suppose that $S_n > S_{n+1}$. Then $\Psi'_{n+1}(S_n - \theta Y_n) \leqslant a^n$ for any θ satisfying $0 \leqslant \theta \leqslant S_n/Y_n$. Since $G_n'(S_n) = 0$, it follows that

$$-1 + \int_{s'}^{1} \frac{1}{\theta} f(\theta)\,d\theta + a \int_{0}^{s'} f(\theta)\,d\theta \geqslant 0$$

for $s' \equiv S_n/Y_n > S_{n+1}/Y_{n+1} = s$. In view of the negative value of (21), the derivative of (11), this is impossible.

For $S_n \leqq S_{n+1}$, $\Psi'_{n+1}(S_n - \theta Y_n) = a^n$ so long as $0 < \theta < S_n/Y_n$. With $s = S_n/Y_n$, the condition $G_n'(S_n) = 0$ is equivalent to equation (11). The uniqueness of the positive solution to (11) has already been discussed.

The rest of the argument is identical to the corresponding demonstration for the case $n = N$. This concludes the proof.

Note that if $R_0 < sY_1$, then $R_n < sY_{n+1}$ for all n. This follows by induction from the fact that $R_{n-1} < sY_n$ implies $M_n^* = sY_n \leqslant sY_{n+1}$, and $\theta > 0$ implies $R_n < M_n^*$, together yielding $R_n < sY_{n+1}$. For our purposes it is essentially irrelevant to consider the unbelievable case where inventory stocks start out so superabundant that no intermediates need be produced $(I_1^* = 0)$. Thus, we assume $R_0 < sY_1$. This justifies our exclusive consideration of $I_n^* + R_{n-1} = sY_n$, $I_n^* + F_n^* = Y_n$ as the relevant planning prescription.

YALE UNIVERSITY