# Plans, Prices, and Corruption: The Soviet Firm Under Partial Centralization, 1930 to 1990

Mark Harrison

Department of Economics
University of Warwick
and
Centre for Russian & East European Studies
University of Birmingham

and

Byung-Yeon Kim

College of Economics
Sogang University

## Acknowledgements

We thank Eugenia Belova, Michael Ellman, Paul R. Gregory, and Hyun Song Shin for comments. We are responsible for remaining errors of fact and interpretation.

Address for correspondence: Harrison, Department of Economics, University of Warwick, Coventry CV4 7AL, email mark.harrison@warwick.ac.uk.

First draft 13 August 2001. This version 11 June 2004.

# Plans, Prices, and Corruption: The Soviet Firm Under Partial Centralization, 1930 to 1990


## Abstract

The level of corruption in an economy is generally thought to influence economic growth adversely. We show that the performance of the Soviet economy was affected not only by the level of corruption but also by its quality, that is, how corrupt incomes were used. In the context of a partially centralized economy, changes in a government control mechanism influenced the quality of corruption and thus economic performance. On the basis of new historical research on the Soviet command system we analyse the choices of a plan-setter and an effort-setter who interact with each other and an external market to determine real output, hidden inflation, and the level and quality of corruption simultaneously. Our results explain rapid Soviet economic growth despite high corruption levels, and why slower economic growth in the 1970s was accompanied by increased privatization of rents.

# Plans, Prices, and Corruption: The Soviet Firm Under Partial Centralization, 1930 to 1990

Bribery and illegal exchange are as old as government and regulation. Moreover, since some regulation has been economically harmful there is often a plausible case to argue that the bribe that enabled government authority to be set aside left everyone better off. But the case for a little corruption disappears if the regulation that was violated would have promoted the common good, or if the government passed a law that was harmful with the intention of enabling its agent to collect the bribe that the violator was ready to pay.

Has corruption in fact helped to lubricate the wheels of the economy, or has it only served to enrich a few at the expense of the many? Corruption has been shown generally to reduce investment and growth (Mauro 1995). It is said to flourish in conditions of authoritarian rule, which hinders accountability and promotes secrecy (Shleifer and Vishny, 1993; Ehrlich and Lui, 1999). However, the variation in growth rates across countries that are similarly corrupt and authoritarian is substantial and exceeds that among democracies (Sah, 1991). This suggests that corruption is unlikely to have the same economic significance everywhere. Economists and historians must evaluate it case by case.

In this paper we consider the case of corruption under Soviet socialism. We look at how corruption arose in a centralized command system in connection with decentralized price setting and the ability of agents to extract side payments for output. There are two main findings. First, we show that Soviet economic performance was affected not only by the level of corruption but also by its quality, that is, how corrupt incomes were used. Second, we show how an instrument of policy, the level of plan "tension," influenced the quality of corruption and thus economic performance. To support these findings we develop a model of a partially centralized economy in which a plan-setter and an effort-setter fix the scale and uses of corrupt side payments simultaneously with effort, output, and inflation. The evidence underlying the model comes from recent investigations in the Russian state archives.

We proceed as follows. Part 1 describes the historical limits on centralization of the Soviet command system. Specifically, Soviet managers exercised discretion over the accounting prices used in planning and prices and side payments that arose in decentralized contracting; they used this discretion to improve the ratio of reward to effort directly, and also to reduce effort indirectly by "siphoning" resources from the retail market. In Part 2 we present a model of the Soviet firm in which there is no corruption and producers allocate effort subject to a plan target and a resource constraint. Part 3 looks at the Soviet firm's use of corruption possibilities to relax its resource constraint through siphoning. In Part 4 we analyse the siphoning process in continuous time and we test an implication using Soviet-era data. At each point we consider the rationale for Soviet managers to engage in corruption and for planners to tolerate or restrain their behaviour. Part 5 concludes.

# 1. The Limits of Centralization

## 1.1. The Soviet Firm as Price Setter

Standard models present the Soviet firm as a price-taker. The source of this stereotype is the official myth of the command system in which administered prices were imposed on producers. Western economists had voiced longstanding empirical concerns (summarized by Harrison, 2000) about the stability of official prices, and especially how the enterprise might exploit product variations to free itself from price controls. This ought to have suggested that the idea of the Soviet firm as price-taker was oversimplified. But such concerns tended to be forgotten when economists offered theoretical generalizations (Nove, 1958; Ames, 1965; Ericson, 1983; 1984; Goldfeld and Quandt, 1988; 1990; Gregory and Stuart, 2001).

According to Kornai (1992: 148) there were too many prices for the price-fixing authority to exercise this function effectively; instead, the authority "merely endorse[d] prices set by the producer" after a process of "vertical bureaucratic bargaining." If producers were free to set prices, how would they set them? Kornai neglected to go on to analyse the producer's pricing decision. This gap in the literature has remained unaddressed except by Shleifer and Vishny (1992), who proposed a model in which the socialist firm faces an incentive to cut both price and output relative to a market equilibrium in order to gain bribes. Given the price, the firm sets output to maximise corrupt side payments. But this approach, which we will call the Shleifer-Vishny conjecture, gives rise to two anomalies. First, in their framework the planner is able to control prices perfectly if she wishes, yet does not control quantities at all. Second, although their conjecture sufficiently explains the phenomena of widespread shortages and corruption in socialist systems, it lacks empirical foundation.

Much detailed information is now available about how Soviet suppliers set prices (Harrison, 1998; Harrison and Simonov, 2000), but even a single example of a firm asking for a lower official price has yet to be found. On the contrary, firms worked continually to secure higher prices. We will review the evidence briefly in relation to plan prices for output, and then the prices actually paid for output in contracts and side payments.

According to the stereotype, the firm in the Soviet command system was managed by non-price regulation of quantities. The Russian archives show us, however, that most targets that mattered in the Soviet economy were denominated in rubles. The most important control figures decided by the Politburo, the annual investment plan and the defence budget, including the plan for military procurements from industry, were fixed in rubles (Harrison 1998, 2001: 99-100: Gregory, 2001: 17-18; Davies, 2001: 64-6). So were most supply quotas that were binding on production ministries and enterprises.

The system of plan prices emerged side by side with centralized supply planning to prevent producers from satisfying ruble quotas by inflation rather than with real output. Supposedly the plan price of each product was fixed for the period of the plan. The anchor for each plan price was the "factory" price officially approved by the centre in some past year, usually on the basis of reported actual direct costs plus an allowance for overheads. The motive for the producer to fulfill the plan then arose because her reward was formed by the degree of fulfillment based on reported real output multiplied by the plan price.

The authorities usually had enough information to prevent producers from adopting an openly inflated plan price for a given product, which would have amounted to open rule-breaking. However, producers could still inflate the overall level of plan prices by changing the assortment profile of output (Harrison, 1998; 2000). Planners controlled both assortment and the setting of plan prices for new, regraded, and unique products with difficulty. When new products were introduced subsequent to the base year for plan prices a benchmark plan price for that product had to be established from scratch, usually on the basis of the actual unit costs that the producer now currently reported. Thus for new products the producer held the initiative in price-setting. In this way product innovation formed the opportunity for producers to exert upward pressure on the average level of plan prices, first by inflating the nominal costs at which new and upgraded products were priced, and second by skewing the assortment towards newer products. This pressure was not costless and to exert it required effort. Still, product innovation enabled producers to achieve an upward drift in plan prices on average that was concealed from the planners at the time although observable in hindsight, and so reduce the effort required to achieve the plan in rubles.

## 1.2. Contract Prices, Liquidity, and Siphoning

Soviet firms used their influence on price-setting not only to fool the planners but also to extract unplanned revenues from purchasers. The documentary evidence on this is plentiful and unambiguous. Stalin’s Politburo took the inflationary propensities of producers for granted; for example, they determined and at times limited the annual investment plan in the light of their understanding that an ambitious plan fixed in rubles could be partly eroded in real terms by higher construction costs: the building industry would respond to a larger budget by raising costs above estimate prices (Gregory, 2001; Davies, 2001).

Records of transactions at lower levels provide further evidence that suppliers’ preferences for product prices were invariably inflationary. In contrast to the firm portrayed by Shleifer and Vishny that gained from a lower price of output, it was typical for Soviet firms to try to push factory prices upward in the direction of the market-clearing level.

The opportunity for inflation arose when suppliers and users negotiated the decentralized contracts that translated aggregate ministerial plans into firm-specific delivery obligations (Kroll, 1986; Harrison and Simonov, 1998; Belova, 2001b; Belova and Gregory, 2002). According to Belova and Gregory it suited planners and suppliers alike that centralized plans were issued in a highly aggregated form; this freed planners from responsibility for issuing specific assignments and gave suppliers freedom of action in deciding how to meet them. In the context of a seller’s market the outcome was a ritual played out each year between ardent suitors (purchasers) and reluctant brides (suppliers) that became known as the annual “contracts campaign.”

Decentralized contracting allowed suppliers to use several stratagems to extract advantage from potential purchasers. As a result purchasers often paid for output above the rate set by officially established factory prices. In the defence industry, for example, Harrison and Simonov (2000) found unauthorized price increases in contracts for established products together with demands for contracts in which prices of new products were specified "provisionally" and subject to review in the light of actual costs, demands for illegal advance payments, and a variety of strategic actions that supported these demands: exaggeration of production costs, withholding of evidence concerning costs, refusal to permit the verification of costs, delaying coming to terms until well into the contract period, and refusal to come to an agreement at all unless concessions were made on the contract price.

Since Kornai (1980) much theorizing about command systems has started from the socialist firm's soft budget constraint. Its financial shortfalls would ultimately be made up from fiscal or quasi-fiscal sources. If so, why did firms invest so much effort in extracting liquidity from purchasers? The answer appears to be that they particularly valued liquidity that was above the plan and, not having a planned use, could be used in a discretionary way.

Why did firms seek to acquire unplanned liquidity? Managers behaved corruptly not just for private embezzlement as Shleifer and Vishny supposed, but also to help the enterprise fulfill its plan (Belova, 2001a). With discretionary cash they could finance side payments to their own workers and to unofficial supply agents (*tolkachi*), and also "siphon" needed but unplanned resources away from the retail market.

The idea of siphoning from the retail market is not as strange as may appear at first sight. Of course the retail market did not supply firms with specialized producer goods such as metal-cutting machine tools, road or rail trucks, steel ingots or nonferrous castings. But many general purpose goods and services that were planned for consumer supply could also be used in production: transport services, fuel and power supplies, information and communication technology, light automobiles and parts, electrical components and fittings, building materials, furniture, and office supplies. These were the commodities over which firms competed with households in the retail market.

Siphoning plays an essential role in the theory of the soft budget constraint. Specifically it solves the riddle posed by Shleifer and Vishny (1992): "In Kornai's model," they wrote, "it is not so much that goods are underpriced, but that the income of the buyers is effectively infinite. This model may be appropriate for some intermediate goods. But households face hard budget constraints, and therefore the systematic shortages of many consumer goods remain a puzzle." In fact a soft budget constraint for firms could result in shortages for consumers if firms' demand for inputs spilled over into retail outlets and siphoned resources intended for final consumption back into intermediate use.

This idea had long been floated in the Sovietological literature (Kaser, 1975; Birman, 1980; Kornai, 1980). However, the existence of siphoning remained conjectural, especially since it required firms to be able to mobilize unauthorized purchasing power. Even if siphoning occurred its scale and significance remained doubtful. If it was significant, it was unclear why the authorities tolerated it.

More recently Qian (1994), Dewatripont and Maskin (1995), and Kim (2002) provided the pieces necessary to solve this puzzle. Dewatripont and Maskin showed how centralization of credit leads to a soft budget constraint in the context of sunk costs and contract renegotiation. Qian showed that when state-owned enterprises compete with households for goods that may be used in both consumption and production, prices below the clearing level may improve efficiency by allowing household consumption to crowd out some bad projects that would otherwise proceed. Finally, Kim showed from postwar archival records that Soviet firms and budgetary organizations did engage in siphoning, that their demand spilled over into the retail market through unauthorized channels in addition to some that were officially sanctioned, and that siphoning made a substantial contribution to repressed inflation and a growing monetary overhang in the Soviet retail market.

Cash was the instrument of siphoning, and siphoning is an established fact, but the full range of mechanisms through which Soviet firms could mobilize cash or freely encash bank deposits are not completely clear. The Soviet financial system was designed to separate monetary flows between two distinct circuits, one for anonymous cash and one for traceable non-cash bank credits. In theory budgetary and state-owned enterprises were allowed to use only non-cash credits in transactions with other organizations. Cash was restricted to a parallel circuit of household transactions. The circuits joined where households were paid cash for labour and used the cash to purchase retail products. In this design the plan steered the real economy, and money tracked the plan, but this sequence depended on strict monitoring of the interface between cash and non-cash money.

From the start, the realities of the financial system departed from this design. Budget constrains were softened and credit conditions became chaotic (Gregory and Tikhonov, 2000). It was widely believed that Soviet enterprises had more cash in aggregate than the plan had authorized (Birman, 1980; Grossman, 1986). This suggests that enterprises gained the above-plan liquidity used for siphoning not only from each other in side payments but also from breaking the central bank's restrictions on cashing bank deposits. The evidence for this is mixed; for different views see Holzman (1960), Granick (1987), Gregory (1990), Lushin (1990), Treml and Alexeev (1994), Woodruff (1999), and Gregory and Tikhonov (2000).

In our own theorizing we assume that the firm cheats the planner through hidden inflation and extracts hidden revenues from the customer, and does not cheat the central bank. This assumption will not significantly affect our conclusions. A model of the macroeconomic equilibrium, however, would need to incorporate the government's budget constraint and the balance sheet of the central bank so as to account for overall liquidity growth.

# 2. Plan-Setting and Effort-Setting

## 2.1. Plan-Setting

In this section we develop a framework in which the firm cheats the planner and behaves opportunistically but not corruptly. At this point it may be asked why the firm did not try to corrupt the planner. The answer appears to be that Stalin and his successors deliberately separated planners from management responsibility and kept their numbers small to ensure their loyalty (Gregory, 1990; Belova and Gregory, 2002).

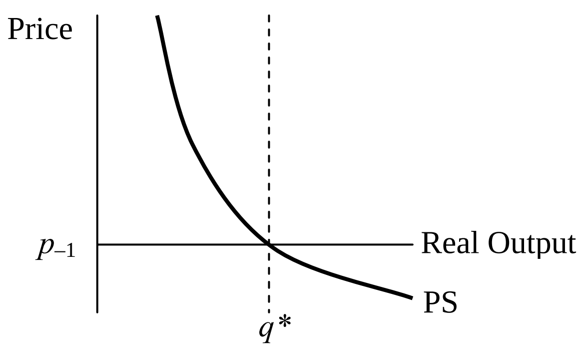
In the Soviet command economy for each period and each firm planners aimed to monitor real output. The firm's target was for its real gross output expressed in plan rubles; this was calculated as a quantity vector $q^*$ multiplied by a vector of official prices $p_{-1}$ that applied in the previous base period. In practice, however, planners could not effectively compel firms to fulfill the plan at base-period prices. When products were changing, planners tried to limit changes in product-unit prices to those that left the level of prices per unit of characteristics unchanged. They could prohibit open inflation, but continuous alterations in the product assortment and product characteristics left planners unable to detect the hidden inflation associated with simulated product innovation. In short, we define hidden inflation in relation to planners, not consumers who knew what prices they faced. Inflation was concealed and the planners were fooled when the firm pushed up product-unit prices faster than the value to the user of the improvement in product characteristics, or when product-unit prices remained constant despite unreported worsening of product quality.

As a result the planners had to be satisfied *ex post* with any real output vector $q$ that, combined with a new current price vector $p$ set by firms and subject to planners' limited scrutiny of price alterations, matched the ruble total set by them *ex ante*. But it is simpler if we treat the firm's average price and quantity as scalars.

$$q = \frac{1}{p} \cdot \left( p_{-1} \cdot q^* \right) \tag{1}$$

This equation is represented in Figure 1 by the PS or plan-setting curve, which was unit-elastic and passed through $p_{-1}, q^*$ when $p = p_{-1}$ and there was no hidden inflation. In general we will only be concerned with that part of the PS curve that lies above the horizontal line marking $p = p_{-1}$. The vertical axis in Figure 1 represents the true level of the official price, including the inflation that the planners failed to detect. The section of the PS curve above the $p_{-1}$ line therefore shows the rate at which producers could trade real output for hidden inflation while continuing to satisfy the plan. Similarly, the horizontal axis measures not the output that was officially recorded but true real output, corrected for the exaggerated claims that producers made for the quality of their products in order to fool the plan-setter.

*Figure 1. The Plan-Setting Curve*




## 2.2. Effort-Setting

From the firm's point of view output was costly, but so was the concealment of inflation, because both required the exertion of effort. Think of the firm's utility as based on its members' wage income $w$ and leisure $\lambda$:

(2) $$u = u(w, \lambda), \;\; u_w > 0 \, , u_\lambda > 0, \;\; u_{ww} < 0, \;\; u_{\lambda\lambda} < 0, \;\; u_{w\lambda} > 0, \;\; u_{\lambda w} > 0$$
.

where $u_w$ and $u_\lambda$ are the first derivatives of *u* with respect to *w* and $\lambda$ respectively, $u_{ww}$ and $u_{\lambda\lambda}$ are the second derivatives of *u* with respect to *w* and $\lambda$ respectively, $u_{w\lambda}$ is the cross derivatives of *u* with respect to *w* and $\lambda$, and $u_{\lambda w}$ is the cross derivative of *u* with respect to $\lambda$ and *w*.[1] The wage could assume two values: a fixed return $w = \overline{w}$ if the firm satisfied the planners' target and otherwise nothing, $w = 0$. This is the classical Soviet incentive scheme in a simplified form (Berliner 1952). We assume that leisure and income interact, so the firm required at least some of both. Berliner first noted the significance of a "quiet life" for the Soviet enterprise. In order to have any income at all the firm had first to satisfy the output target, and then it could also maximize leisure. The firm maximized leisure subject to a resource constraint that we shall treat initially as hard:

(3) $$1 - e - i - \lambda = 0$$

where *e* is productive effort, *i* is the effort required to simulate innovation and so conceal inflation, and the firm's initial time endowment is normalized to 1.

Finally, there were two technologies, one for the production of output *a* and one for the concealment of inflation *x*. Production required effort to be combined with capital *k*, but for now we will treat capital as an exogenous resource. Inflation concealment required effort alone. In both activities, returns to effort diminished:

(4) $$q = a(\overline{k}, e), \quad a_e > 0, \quad a_{ee} < 0$$

(5) $$\frac{p - p_{-1}}{p_{-1}} = x(i), \quad x_i > 0, \quad x_{ii} < 0$$

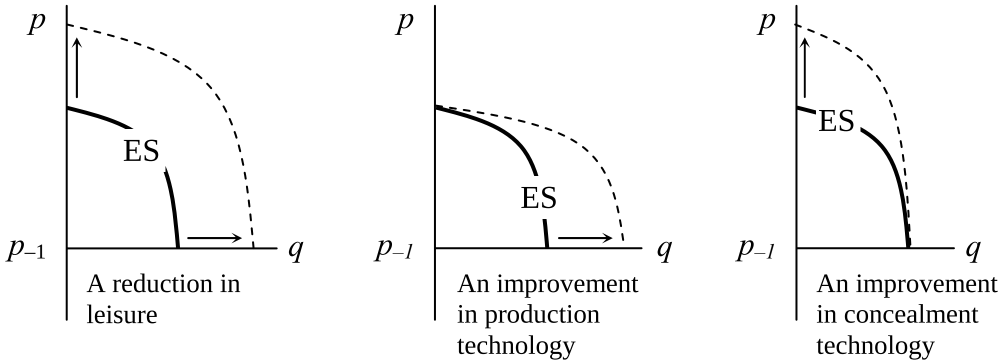
Then, for any given level of overall effort that it chose to set, the firm faced a feasible set of combinations of real output and hidden inflation that was concave to the origin (**proposition 1**: for propositions, proofs, and a list of symbols see Appendix 1).

We call this the ES or effort-setting curve. Figure 2 shows that for each firm and for every production and concealment technology there was a family of ES curves, one for each feasible level of effort. As the firm set its effort level higher, the ES curve moved outwards in all directions (**proposition 2**). An improvement in production technology shifted the ES family to the right, raising real output for given effort and hidden inflation. An improvement in the concealment technology shifted the ES family upward, raising hidden inflation for given effort and real output (**proposition 3**).

---

[1] The same definitions of first and second derivatives apply to other equations.

*Figure 2. The Effort-Setting Curve*




## 2.3. The Firm's Equilibrium

Facing a given PS the firm's problem was to allocate effort between production and inflation concealment so that the plan was fulfilled and effort was set at a minimum. In Figure 3 this point is found at A where the plan-setter's line is tangential to the lowest available effort-setter's line, marked by $\lambda_0$, and it sets the equilibrium values of real output and hidden inflation simultaneously. The firm could also have satisfied the planner at B with more real output and less hidden inflation, or at C with less real output and more hidden inflation, but both of these would have cost the firm more effort, $\lambda_1$.

*Figure 3. The Firm's Equilibrium*

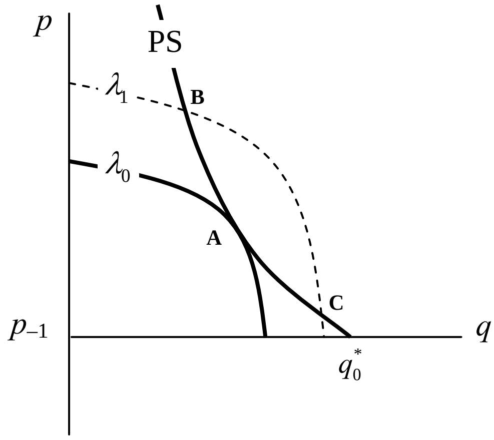
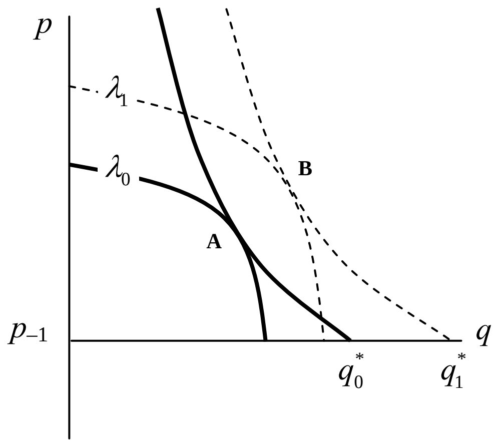



Figure 4 illustrates what happens when the planner increases tension. Other things being equal, an increase in $q*$ right-shifts PS. The firm is forced to move to the higher ES curve marked by $\lambda_1$. At the new equilibrium B there is less leisure and more real output, but diminishing returns to productive effort make it inefficient for the firm to meet the higher plan only by increasing real output. Some of the extra effort goes into fooling the planners. Therefore real output rises by less than $q*$ and there is also more hidden inflation.

*Figure 4. A Rise in Plan Tension*




Simple extensions of the model suggest conditions under which the command system may break down. These conditions include both excessive liberalism and excessive harshness. First, suppose the plan was set at a level that could not be fulfilled with any combination of real output and hidden inflation given the firm's hard resource constraint. Producers would then prefer zero effort, zero reward, and zero utility to the loss resulting from trying and failing to fulfill the plan; there would be no hidden inflation, but output would collapse to the origin. One is reminded of the turmoil induced by Stalin's Great Breakthrough and Mao's Great Leap Forward. Second, if prices were fully liberalized within the command system producers would use their new discretion to climb the PS curve, raising prices without limit and cutting real output to a minimum. Again, one thinks of the Soviet economic collapse at the end of the 1980s (Harrison, 2002).

# 3. Corruption and Siphoning

## 3.1. A Single Transaction

So far we have assumed that the enterprise was subject to a hard resource constraint. In fact, while resources for production and consumption were constrained in total for the economy as a whole, the resource constraint on the individual firm could be softened by collecting corrupt revenues and recycling them through siphoning, and so adding outside resources intended for consumption to the production inputs officially allocated to the enterprise. Our attempt to capture the outcome will resemble the Shleifer-Vishny story of the bribe in some respects, but in our initial model the bribe is collected in order to fulfill the plan, not to line pockets. The result is that, with plan-setting unchanged, real output rises while both hidden inflation and insider effort fall; however, the inflation experienced by consumers also rises.

Subsequently we will examine the implications of bribe-taking for personal enrichment, or embezzlement. However, unlike Vishny and Shleifer we do not regard embezzlement as the general case. We will draw a distinction between corruption and disloyalty. All managers were potentially *corrupt*, but only *disloyal* managers used the proceeds for personal enrichment; loyal managers used them to fulfill the plan. Such loyalty was rational. In the vertical administrative hierarchy managers were expected to invest in loyalty both upwards and downwards. Managers showed loyalty to their superiors for the sake of their careers: to get promotion, they needed a record of plan fulfillment. But this record also relied on meeting the needs of inferiors and maintaining their cooperation, so managers also had to show loyalty to the workforce. Consistent disloyalty in either direction, upward or downward, prejudiced career aspirations. Since these loyalties could come into conflict successful management involved complex balancing. Of course disloyal corruption that was for personal enrichment also brought a return, but the value of the return was less than might be thought since cash was often hard to spend on personal consumption; this was both because of the seller's market and also because of the need to avoid attracting attention.

Belova (2001a) reports what she calls the "honest" manager's dilemma, but in this system no one was completely honest, so we will call it the dilemma of the *loyal* manager. To fulfill the plan, the manager had to compete for scarce supplies that the planners forgot, or promised but did not deliver. To compete, she had to pay. Payments went to sideline suppliers of deficit commodities, unofficial agents (*tolkachi*) for purchasing services, and workers for the extra effort required to make bricks without straw. We will analyse siphoning: so as to augment their own productive stocks firms entered the retail market as purchasers of commodities intended for sale to households for personal consumption. For this purpose the loyal manager had to acquire discretionary liquidity by creating revenues that were hidden from the planners. She got the liquidity from her own purchasers by securing an above-list price, or advance payment, or bribe, but, in distinction from the Shleifer-Vishny story, the bribe was credited to the firm's account.

First, rewrite the firm's production function from equation (4), adding $\Delta k$ to represent the availability of an outside resource that substitutes perfectly for the firm's inside stocks of fixed and working capital:

$$q = a(\overline{k} + \Delta k, e),\ \Delta k > 0,\ a_e > 0,\ a_{\Delta k} > 0,\ a_{ee} < 0,\ a_{\Delta k \Delta k} < 0,\ a_{e\Delta k} < 0,\ a_{\Delta k e} < 0 \tag{6}$$

.

In order to get hold of $\Delta k$ the firm requires hidden revenue for discretionary use in the secondary market where commodities are traded that are of potential intermediate use. This hidden revenue must be over and above the liquidity officially allocated in the firm's financial plan. The firm extracts this hidden revenue in advance from consumers in its own product market: the hidden revenue gathered in the previous period, $b_{-1}$, is used to add to the firm's real resources in the current period. We model this as follows: there is a market-clearing price $\hat{p}$ of output that exceeds the official price $p$ at all relevant levels of output. The share of goods sold at the market-clearing price, $\hat{p}$, is $r$ and the remaining goods are sold at the official price, $p$. The existence of both regulated and market-clearing prices in shortage economies is a well established fact that can be rationalized in economic theory by reference to consumers' heterogeneous valuations of time versus money (Kim, 1997).

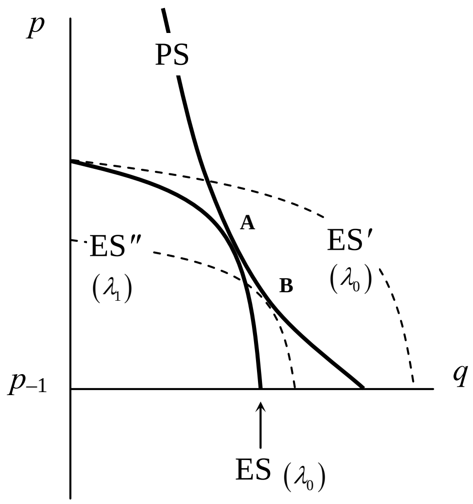
Suppose the firm can capture some proportion of the gap between the regulated and market-clearing price in the form of an advance payment, side payment, or markup over its officially listed product price without being detected; for simplicity we can set this proportion at 100 per cent without affecting basic results. The firm uses its hidden revenue as discretionary purchasing power in the secondary market. The effect on the firm's resources is the outcome of a siphoning technology $s$ in which the input is hidden revenue but, since the secondary market too is a seller's market, there are diminishing returns:

(7) $$\Delta k = s(b_{-1}),\ \ s_{b_{-1}} > 0,\ \ s_{b_{-1}b_{-1}} < 0$$

(8) $$b_{-1} = r \cdot q_{-1} \cdot (\hat{p}_{-1} - p_{-1}),\quad 1 > r > 0,\quad q > 0,\quad \hat{p} > p$$

How will siphoning affect the ES curve? The result will be a new ES′ curve that is still concave (**proposition 4**) but right-shifted by comparison with the no-siphoning case (**proposition 5**). For a baseline illustration, consider how the structure of the firm's incentives is changed in what we will call the "single-transaction" case. The single transaction begins in period –1 when the firm sells a commodity in return for a side-payment and recycles it through siphoning to acquire $\Delta k$. The same transaction ends in period 0 when the firm uses $\Delta k$ to satisfy the plan-setter and augment its utility, but the firm makes no provision for repeated siphoning of resources to be used in period 1.Thus the single transaction takes two periods to complete, but the only decision we analyse is the firm's optimization within the current period.

*Figure 5. Siphoning: A Single Transaction*




This case is illustrated in Figure 5. The firm is initially subject to a hard resource constraint. Given plan-setting, the firm is in equilibrium with the solid effort-setting curve ES and leisure $\lambda_0$ at point A. Softening its resource constraint through siphoning enables the firm to produce more for given hidden inflation and given effort. The dashed ES′ curve illustrates the production-augmenting effect of siphoning with effort held at $\lambda_0$. Since ES′ lies partly outside the firm's PS curve, it is not efficient for the firm to hold effort at this level. As long as plan-setting remains unchanged, the firm will prefer to cut effort to $\lambda_1$, yielding an ES″ line that, with augmented productive capital, is skewed to the right by comparison with the no-siphoning ES curve and touches PS at B, a point showing more real output and less hidden inflation.

In short, behaving corruptly to soften her resource constraint helped the loyal manager to fulfill the plan with less effort, less hidden inflation, and more true real output than would have been possible otherwise. This result suggests that not only the firm gained; there was also less fooling of the planners, for whom there was a clear rationale to turn a blind eye to the rule-breaking involved since in the outcome the plan was fulfilled more honestly.

The welfare implications are ambiguous, however. Consumers are better off in the current period in terms of real consumption because real output has increased; as long as production finishes at the end of the current period we can assume that all goods produced currently are consumed. But it is not clear that this period's consumption will have risen by enough to compensate for the consumption lost in the last period by being recycled back into production. Equations (6) to (8) show that this depends on the efficiency with which the firm captures hidden revenues and converts them into additional inputs, and on the firm's initial endowment; a firm that starts capital-poor will make more productive use of siphoning.

For consumers subject to a cash constraint the effect of siphoning on prices must also be taken into account. Consumers pay less for supplies in the current period because of less hidden inflation but in the previous period some of them paid more for supplies because of a combination of higher free market prices and covert side payments.

Next, we show that the rise in plan tension shown in Figure 4 can induce enterprises to use siphoned resources for higher output rather than effort reduction; thus, consumer welfare was more likely to improve when siphoning was combined with a rise in plan tension.

## 3.2. Loyal and Disloyal Managers

In this economy all managers were self-interested, and all managers were corruptible, but the manager we have described so far remained loyal. By accepting illegal side-payments and disbursing them via the siphoning mechanism, the manager fulfilled the plan and at the same time reduced the efforts of the workforce. This manager was therefore loyal simultaneously to superiors and inferiors.

Were all managers loyal? Above we suggested some reasons why managers, however corrupt, might choose to remain loyal. Nonetheless disloyal managers existed. If disloyal, they pocketed bribes for personal enrichment and failed to share gains with either planners or workers. They still aimed to fulfill the plan but they were no longer minimizing the workers' or their own efforts. It is widely held that disloyal corruption increased through time and contributed to the decay of the Soviet command system under Brezhnev (Grossman, 1998). Therefore the implications of disloyalty deserve brief examination.

A manager might choose to be either completely or partly disloyal. A *completely disloyal* manager embezzled all side payments for personal enrichment, so there would have been no siphoning and the firm's resource constraint would remain hard. The fact of widespread siphoning implies that complete disloyalty was not typical. We suggest that in practice the scope for managerial disloyalty was limited by workers' and planners' responses.

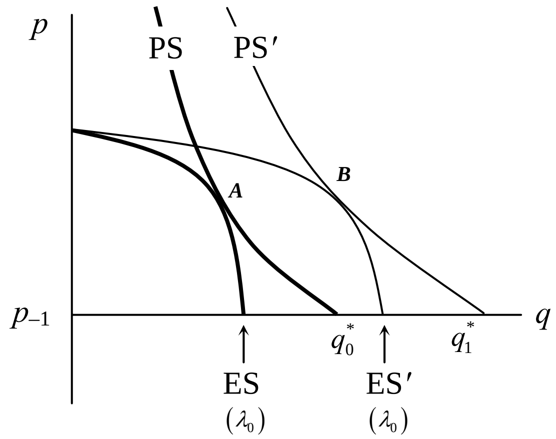
To support this we specify the distribution of information about disloyal behaviour by managers as follows. Suppose that information about particular disloyal acts was only available to other agents within the firm. Above the firm, there was information only about the general incidence of corruption possibilities. If the manager was disloyal the workers knew, but the planners could only suspect.

With respect to the workforce the manager with a disloyal propensity faced the following problem. She could sufficiently maintain her reputation upward by fulfilling the plan. But to fulfill the plan the manager needed the cooperation of the workers, who would have knowledge of her disloyalty. The workers might threaten to withhold effort, or alternatively to disclose her wrong-doing, to induce her to share the gains from corruption. In principle this sharing might be done in two ways. First, the corrupt manager might distribute part of the bribe income directly to the workforce in cash, but on our understanding of the context and evidence this was detectable, therefore dangerous, and seldom done. Instead, managers bought cooperation by transforming cash rents into additional resources through siphoning and sharing the gain with the workers in the form of leisure. As a result, for a given plan insiders' effort fell. In this form rent-sharing was less visible to higher-level audit or even looked "good." In short, complete disloyalty was not feasible: *self-interested managers were always at least partly loyal*.

Consider the problem of the same disloyal manager in relation to plan-setters. Plan-setters did not know firms' intrinsic capacity, or the extent of disloyalty, or the extent to which bribes were recycled into siphoning. They knew there was corruption, that siphoning might occur as a result, that siphoning enlarged firms' capacity, and that even disloyal managers would rationally choose to recycle some bribes into siphoning. They knew therefore that when siphoning arose the effort level and the utilization of capacity would tend to fall with an unchanged plan; this case was illustrated in Figure 5. Thus they would rationally respond to information about increased corruption possibilities by tautening the plan.

Figure 6 shows a possible outcome. The firm's capacity gain from siphoning, controlling for effort, is the shift from ES to ES′. The increase in plan tension is the shift from $q_0^*$ to $q_1^*$ and from PS to PS′. In the case of a corrupt but loyal manager this increase exactly captures the firm's additional resources, moving its equilibrium from A to B. At B, effort is the same as at A. Managers' loyalty has resulted in the gain being shared with the planners through the risen in real output, but real output has risen less than in proportion to the rise in plan tension and there is also higher hidden inflation.

*Figure 6. A Rise in Plan Tension With Siphoning*



The disloyal manager wishes to pocket the gain herself. But she still has to conserve her reputation with plan-setters by fulfilling the plan. Since the planners' knowledge of corruption possibilities has led them to set a higher plan, the disloyal manager must now get the workers to work harder. For this she needs the workers' cooperation, and this forces her to share the proceeds of corruption with the workers by engaging in siphoning to enlarge capacity and limit effort. In short, *when managers were corrupt raising plan tension was a mechanism to limit disloyalty.* But the converse was also the case: *reducing plan tension could promote embezzlement by corrupt managers*. This result establishes a clear mechanism that links the spread of disloyal corruption through the Soviet economy in the 1970s with the simultaneous reduction of growth targets (on corruption see Grossman, 1977 and 1979, and on growth targets and growth Schroeder, 1985).

# 4. Repeated Siphoning

The Shleifer-Vishny model prompts the question whether a corrupt supplier can gain by restricting output. To address this question we go beyond the case of a single transaction. In the single-transaction case the value of side payments is fixed beforehand. Effort is not optimized to supply the market in such a way as to secure further side payments as a basis for repeated siphoning. Side-payments depend on the gap between the market-clearing price and the official price. When siphoning is continuous this gap can be determined simultaneously with the effort-setting decision since market-clearing and official prices both decline, but at differing rates, as real output rises. The official price varies inversely with output along the PS curve, the latter being unit price-elastic. The market-clearing price varies inversely with output along the demand curve as the balance in the market shifts from seller to buyer. Therefore, rewrite equations (7) and (8) in continuous time:

(9) $$\Delta k = s(b),\ s_b > 0,\ s_{bb} < 0\,,$$

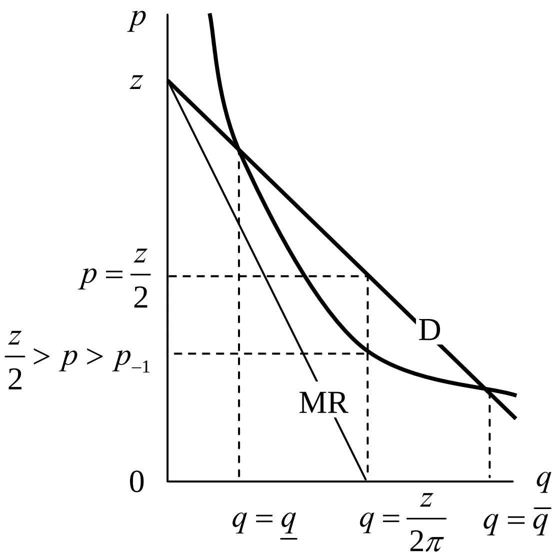
(10) $$b = r \cdot q \cdot (\hat{p} - p),\quad 1 > r > 0,\quad q > 0,\quad \hat{p} > p\,;$$

and add a market demand curve, linear for the sake of illustration:

(11) $$\hat{p} = z - \pi \cdot q,\quad z, \pi > 0\,.$$

Consider how the firm's hidden revenues change as real output increases. We observe that the firm's hidden revenues depend on the relative elasticities of the PS and market demand curves. The PS curve, which could also be termed the planner's demand curve, determines open revenues but is unit-elastic: the firm must achieve a fixed total planned revenue, which in this context becomes a lump-sum tax. Assume for simplicity that the firm succeeds in extracting all the side payments that the market will bear; then, the market demand curve determines the firm's total revenue. It is downward-sloping and linear so its elasticity falls as output rises. Where the elasticity of the market demand curve equals unity the firm's total revenue is maximized and, since its planned revenues are a lump sum, its net hidden revenues are also maximized.

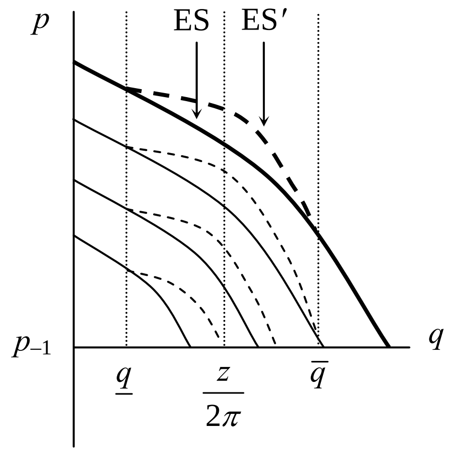
*Figure 7. Hidden Revenues*




The siphoning firm operates in the "shortage" region of Figure 7, in the region between $\underline{q}$ and $\overline{q}$ where the market-clearing price exceeds the official price; it is also a necessary condition that both prices exceed the baseline plan price $p_{-1}$. In this region, assuming that the firm is able to extract the market price on its total output, the gap between the PS and D (market-demand) curves represents the firm's hidden revenue per unit of output. It maximizes its total hidden revenues midway between $\underline{q}$ and $\overline{q}$, which is also the point where marginal revenue becomes zero and the marginal revenue curve, drawn with slope $-\frac{1}{2\pi}$, meets the quantity axis (drawn in this figure at $p = 0$, not $p = p_{-1}$). On either side of this point hidden revenues diminish, falling to zero at $\underline{q}$ and $\overline{q}$.

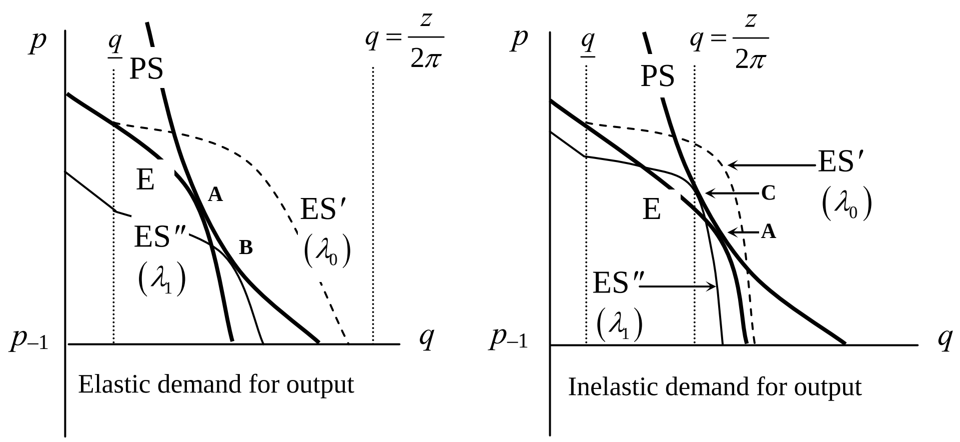
How is the ES curve affected when side payments are endogenous? The result is a blister on the ES curve in the region between $\underline{q}$ and $\overline{q}$. In the single-transaction case (Figure 5) the whole ES curve was right-shifted in favour of higher real output. When siphoning is repeated the displacement of the ES surface is limited to the region already identified and is upward, reaching its maximum extent at the level of output where hidden revenues are maximized. This drift is determined by the PS and D curves alone, and there is therefore an identical blister on each ES curve in the family of curves as shown in Figure 8.

*Figure 8. The* ES′ *Curve With Repeated Siphoning*




How does repeated siphoning influence the PS-ES equilibrium? For simplicity rewrite equation (9) such that the relationship between hidden revenues and resources siphoned becomes linear with constant returns. The ES′ curve still has a negative slope (**proposition 6**). Now siphoning is maximized where hidden revenues are maximized. Thus, over the output range where the linear market demand curve is *more* elastic than the PS curve and hidden revenues are *increasing*, the siphoning blister makes the ES′ curve *flatter*, and conversely. Under continuous siphoning, therefore, the downward slope of the ES′ curve varies directly with that of the market demand curve (**proposition 7**).

*Figure 9. Repeated Siphoning*




If ES′ is made flatter then the siphoning equilibrium will tend to shift in favour of higher output; conversely, if steeper the siphoning equilibrium will tend to shift in favour of higher hidden inflation (**proposition 8**). The two cases are illustrated in Figure 9. In each case point A marks the no-siphoning equilibrium. The left hand panel shows the case where, with a low value of $\pi$, the market demand curve is still elastic at the no-siphoning equilibrium; then for given output the ES′ curve will become flatter, and the siphoning equilibrium will tend to shift in favour of higher output. Intuitively, the firm that faces elastic market demand loses hidden revenue by restricting output. Instead, up to a point dictated by diminishing returns to effort and liquidity the firm produces extra output, takes the extra hidden revenue, and uses it to buy extra inputs. For insider effort set at $(1-\lambda_0)$ the ES curve expands to ES′. This effort is above the efficient level, so the firm cuts effort to the lower ES″ where effort is $(1-\lambda_1)$ and $\lambda_1 > \lambda_0$. The outcome is an equilibrium at B with less hidden inflation and more real output.

The welfare implications of continuous siphoning for consumers in the case of a relatively flat market demand curve are again ambiguous. Consumers can enjoy higher consumption, as in the single transaction case, but only if the volume of true real output increases by enough to compensate for the loss of consumption due to continuous siphoning. As for the aggregate price level that consumers face, this can go up or down depending upon three factors: the decrease in hidden inflation compared to the no-siphoning case, the excess of the market-clearing price over the official price, and the proportion of consumers paying the market-clearing price.

The right hand panel of Figure 9 shows the converse case. With a higher value of $\pi$ the market demand curve is already inelastic at the no-siphoning equilibrium; then for given output the ES′ curve will be made steeper, and the siphoning equilibrium will tend to shift in favour of higher concealed inflation. The firm facing inelastic demand can raise hidden revenue from side payments by restricting output. This case is more consonant with the stylized facts of a shortage economy proposed by Shleifer and Vishny. But in our case the firm, while restricting output to increase hidden revenues, must still satisfy the planners. It does so by diverting the effort released by external resources siphoned from the secondary market to hiding inflation rather than producing output. The outcome at C is more hidden inflation and less real output.

Thus when market demand is inelastic the implications of continuous siphoning for consumer welfare are clearly negative: consumers become worse off in terms of both consumption and the price level. Resources siphoned away from consumption result in lower output; not only does the official price increase but, in addition, some consumers must also pay the still higher market-clearing price.

This aspect of our model yields a testable prediction: other things being equal, hidden inflation should be inversely correlated with the price elasticity of market demand at the firm level. Specifically, we can rewrite equation (4) for the case of repeated siphoning in the form:

$$\left(\frac{p-p_{-1}}{p}\right)_j = p\left(\eta_j, X_j\right), \quad p_{\eta_j} < 0, \; p_{X_j} < 0 \tag{12}$$

where the dependent variable is the inflation concealed by the *j*th firm, $\eta_j$ is the elasticity of market demand for the firm’s product, and $X_j$ is a vector of industry-level fixed effects that determine the firm’s private effort costs of operating the concealment technology.

Here are three problems: to observe inflation that has been hidden, to identify the elasticity of the firm’s market demand curve, and to control for fixed effects across industries. In Table 1 we measure hidden inflation across branches of industry over the period 1951 to 1987 as the excess of Soviet volume index numbers over CIA estimates that were designed to strip hidden inflation from the Soviet official figures. According to the table, hidden inflation was highest in machinery and lowest in fuel.

*Table 1. Real Growth and Hidden Inflation in Soviet Industry by Branch, 1951 to 1987*

| | Real output growth, per cent per year | | Hidden inflation, per cent per year | Price elasticity of firm's market demand | Inflation con-cealment cost |
|---|---|---|---|---|---|
| | CIA | Soviet | | | |
| | (1) | (2) | (3) | (4) | (5) |
| Machinery | 5.6 | 11.2 | 5.6 | low | low |
| Light industry | 4.3 | 5.4 | 1.1 | high | low |
| Chemicals | 7.8 | 10.6 | 2.8 | low | medium |
| Wood, pulp, and paper | 3.2 | 5.2 | 2.0 | medium | medium |
| Construction materials | 7.0 | 9.0 | 2.0 | medium | medium |
| Food industry | 5.1 | 5.8 | 0.7 | high | medium |
| Ferrous metals | 5.2 | 6.3 | 1.1 | low | high |
| Electric power | 8.1 | 8.9 | 0.8 | low | high |
| Fuels | 5.5 | 5.6 | 0.1 | medium | high |

*Sources and Notes.* Columns 1and 2 are from CIA (1990), 37. Column 3 is column 2 minus column 1. The CIA methodology has been exhaustively described in many publications. Although not above criticism it remains the best attempt at an extraordinarily difficult and complex task. On western attempts to measure hidden inflation over the period before 1955 see Harrison (2000). Column 4 is based on Appendix 2, Table 2.1, column 1, with 0 = "low," 1 = "medium," and 2 = "high.". Column 5 provides estimates of the relative opportunity for "simulated" innovation and direct falsification of output returns in different branches of Soviet industry. These estimates are based on our reading of a number of classic works including Hodgman (1954), Grossman (1960), Kaplan and Moorsteen (1960), Moorsteen (1962), Nutter (1962), and CIA (1979) in addition to more recent studies by Khanin (1991), Harrison (1996), Kudrov (1997), and Harrison (2000). These studies support the rankings offered in the table rather than cardinal measures.

For the elasticity of demand and the effort cost of inflation concealment across industries we provide guesswork that should be considered subjective, although supported in various ways; sources and further detail are provided in Appendix 2. We observe the factors underlying the firm's elasticity of demand in three different ways. First, the structure of each industry determined the buyer's ability to choose among alternative suppliers within the industry; the buyer could shop around among official and unofficial suppliers to a varying extent, and the opportunity for this arose in the contract stage of plan implementation that has already been described. Second, the buyer had more or less scope to substitute within each industry's range of products. Third, the frustrated buyer had varying choices among the substitutes for each industry's products available from other industries, unofficial markets in the "second" economy, or the buyer's own in-house sideline production. We take the median of the three resulting indexes to estimate the relative elasticity of the demand curve facing the individual firm in each industry; in Table 1 we classify it as low, medium or high. We conclude that price elasticity was high in the light and food industries, low in machinery, chemicals, and electric power, and medium in other branches.

As for the effort cost of inflation concealment, we set it low in machinery where rapid product innovation was intrinsic to technological change throughout the economy, and in light industry where consumers placed a premium on novelty, high in electric power, ferrous metals, and fuels where standardization and product homogeneity were easily monitored and at a premium, and at an intermediate level in other branches.

The table shows that, controlling for concealment costs, hidden inflation increased as firm-level demand became more inelastic across industries.[2] This appears to substantiate our theorizing.

Our finding suggests that when siphoning was repeated successive increases in plan tension would no longer necessarily shift the firm's equilibrium towards higher levels of real output. This is because the firm's point of maximum hidden revenues was fixed by the market demand curve. In terms of Figure 9, other things being equal, a rise in plan tension shifted the firm's no-siphoning equilibrium to the right, and at the same time the probability increased that the no-siphoning equilibrium would drift to the right of the point where the firm's hidden revenues were maximized. At this point the firm might rationally prefer to curtail any further increase in real output that would drive down hidden revenues and reduce siphoned resources. This placed an ultimate limit on the ability of the planner to raise plan tension so as to maintain productive effort and raise real output when siphoning was repeated.

Other things being equal, however, an increase in plan tension would reduce the scope for both corruption and siphoning. This was evident in Figure 7: the firm may extract a net surplus of hidden revenues only when market demand exceeds planned revenues. An increase in the firm's planned output reduces the gap between the two and so limits the firm's corruption possibilities. In short, when siphoning was repeated and firms responded to the resulting possibilities with adverse effects on real output, regardless of whether managers were loyal or disloyal, *tautening the plan was a mechanism to limit corruption and corruption-oriented behaviour.* Conversely, reducing plan tension increased the scope for both corruption and disloyalty.

---

[2] Consider the following, purely for illustration. In the table, set "low," "medium," and "high" equal to 0, 1, and 2 respectively, as in appendix 2, table 2.1; then hidden inflation may be regressed on demand elasticity and concealment costs. With an adjusted $R^2$ of 0.86, despite the small sample size both slope coefficients are negatively signed and significant at the 0.5% level. Controlling for concealment costs, an increase in the imputed elasticity of demand from "low" to "medium," or from "medium" to "high," reduces hidden inflation by 1.47 percentage points. For given demand elasticity the same increase in imputed concealment costs reduces hidden inflation by 1.86 percentage points. However, the independent-variable data are far from robust.

# 5. Conclusions

We have analysed the Soviet firm as a price-setter that interacted with a plan-setter and an external market to determine real output, hidden inflation, corruption, and siphoning simultaneously. In the case of a single siphoning transaction we have found a clear rationale for Soviet planners to tolerate corruption and siphoning. Managers allocated resources that were gained corruptly to produce more real output with less hidden inflation and fulfill the plan more honestly as a result. Self-interested managers, however corrupt, always remained at least partly loyal to the goals of planners and workers. Moreover, by increasing plan tension planners could induce disloyal managers to recycle more hidden revenues into \production. Thus, tautening the plan limited managerial disloyalty. The implications for the welfare of Soviet consumers are ambiguous, however.

We have qualified these results for the case of siphoning repeated in continuous time. In this case higher real output was not the invariable consequence. When siphoning was continuous the consequences for the welfare of consumers were less likely to be positive. But raising plan tension appears to have remained an effective mechanism for limiting the scope for corruption and adverse corruption-oriented behaviour.

Our findings also reflect on the relationship between Soviet corruption and shortage. We are far from convinced that the shortage economy was created with the intention of distributing bribes to individuals. Rather, the shortage economy had a by-product, corruption, that planners tolerated for their own purposes. Corruption did create scope for disloyal agents to line their pockets, but it could also help to align the objectives of suppliers with those of planners. Planners could influence the scope for both corruption and disloyalty by manipulating the degree of plan tension; it seems likely that reductions in plan tension in the 1970s contributed to the privatization of rents in the Soviet economy.

The relationship between corruption and growth that we have found in the Soviet economy may apply more widely in countries where developmental states impose partial centralization. In such economies firms pursue multiple objectives. When this happens they are subjected to regulation. Regulated enterprises can then choose to allocate resources between adding value, seeking rents, and augmenting leisure. Because the economy is only partly centralized, they can also attract external resources by means that qualify as corrupt. These additional resources may in turn be applied to the purposes of personal enrichment or to those of the enterprise itself. We suggest that, controlling for the level of corruption, economic performance depends on how corrupt incomes are used. An authoritarian regime that tolerates corruption conditionally upon economic performance can reduce the scope for private enrichment and maintain growth by encouraging the siphoning of resources back into production. This may explain why some countries under authoritarian regimes have experienced rapid economic growth in spite of significant corruption.

# Appendix 1. The Effort-Setting Curve

## Symbols

| | |
|---|---|
| $a$ | production technology |
| $b$ | firm's hidden revenue |
| $e$ | effort in production |
| $i$ | effort in concealment of inflation |
| $k$ | capital |
| $\lambda$ | leisure |
| $p$ | official price |
| $\hat{p}$ | market-clearing price |
| $\pi$ | (absolute) slope of the market demand curve for output |
| $q$ | real output |
| $r$ | share of output sold at the market-clearing price |
| $t$ | plan target |
| $x$ | inflation concealment technology |

## Plan-Setting and Effort-Setting

Consider a firm that operates subject to two constraints. The first constraint is the PS or plan-setting line: its real output at official prices must match its target in rubles at plan prices, $p_{-1} \cdot q^*$, deflated by the current official price level:

(A1) $$q = \frac{1}{p} \cdot \left( p_{-1} \cdot q^* \right).$$

Second is a resource constraint, where the firm's time endowment is normalized to 1:

(A2) $$1 - e - i - \lambda = 0.$$

The firm has two activities. Production requires capital and effort. To begin with only effort is variable; effort is also subject to diminishing returns:

(A3) $$q = a \cdot \bar{k}^{1-\alpha} \cdot e^{\alpha}, \quad 0 < \alpha < 1.$$

The concealment of inflation in the price at which the firm's plan target is to be fulfilled requires only effort, again subject to diminishing returns:

(A4) $$\frac{p - p_{-1}}{p_{-1}} = x \cdot i^{\beta}, \quad 0 < \beta < 1,$$

where $p$ is concealed from the plan-setter.

For reasons given in the text the firm's problem, given the plan-setter's line, is reduced to allocating effort between production and inflation concealment so as to fulfill the plan with minimum overall effort. This point is found where PS, the plan-setting line (equation A1), is tangential to the lowest effort-setting line or ES (from A2, A3, and A4):

(A5) $$q = a \cdot \bar{k}^{1-\alpha} \cdot \left[1-\lambda-\left(\frac{p-p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right]^{\alpha}.$$

## The ES Curve

**Proposition 1.** *The tangential line of the* ES *curve is negative. In addition, the* ES *curve is strictly concave: the marginal rate of substitution decreases as p increases.*

*Proof:*

The partial derivative of $q$ with respect to $p$ is less than zero:

$$\frac{\partial q}{\partial p} = a \cdot \bar{k}^{1-\alpha} \cdot \alpha \left[1-\lambda-\left(\frac{p-p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right]^{\alpha-1} \cdot \left(-\frac{1}{\beta}\right)\left(\frac{p-p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}-1}\left(\frac{1}{x \cdot p_{-1}}\right).$$

Set $A = \left[1-\lambda-\left(\frac{p-p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right] > 0$, and $B = \left(\frac{p-p_{-1}}{x \cdot p_{-1}}\right) > 0$.

Then $\frac{\partial q}{\partial p} = -\frac{a \cdot \bar{k}^{1-\alpha} \cdot \alpha}{x \cdot \beta \cdot p_{-1}} A^{\alpha-1} \cdot B^{\frac{1-\beta}{\beta}}$

and $\frac{\partial q}{\partial p} < 0$.

Moreover, the second derivative of $q$ with respect to $p$ is also negative:

$$\frac{\partial^2 q}{\partial p^2} = -\frac{a \cdot \bar{k}^{1-\alpha} \cdot \alpha \cdot (1-\alpha)}{x \cdot \beta \cdot p_{-1}} \cdot A^{\alpha-2} \cdot B^{\frac{1-\beta}{\beta}} \cdot \left(\frac{1}{x \cdot p_{-1}}\right)^2 \cdot \left(\frac{p-p_{-1}}{\beta}\right) - \frac{a \cdot \bar{k}^{1-\alpha} \cdot \alpha}{x \cdot \beta \cdot p_{-1}} \cdot A^{\alpha-1} B^{\frac{1-2\beta}{\beta}} \left(\frac{1-\beta}{\beta}\right)\left(\frac{1}{x \cdot p_{-1}}\right)$$

,

where $0 < \alpha < 1$ and $0 < \beta < 1$. Thus $\frac{\partial^2 q}{\partial p^2} < 0$, suggesting that the ES curve is strictly concave to the origin.

**Proposition 2.** *A reduction in leisure expands the* ES *curve; the new curve lies outside the old curve at all points.*

*Proof:*

This proposition is investigated through the effects on the $p$- and $q$-intercepts of the ES curve arising from a decline in the value of $\lambda$. The $p$-intercept is found by setting $q = 0$ in equation A5:

(A6) $$p = p_{-1} \cdot \left[x \cdot (1-\lambda)^{\beta} + 1\right].$$

Similarly the $q$-intercept is found from equation A5 by setting $p = p_{-1}$:

(A7) $$q = a \cdot \bar{k}^{1-\alpha} \cdot (1-\lambda)^{\alpha}.$$

Other things being equal, when $\lambda$ declines, both the *p*- and *q*-intercepts increase: the ES curve expands in all directions.

**Proposition 3.** *An improvement in production (concealment) technology shifts the* ES *curve toward output (concealed inflation).*

*Proof:*

Consider equations A6 and A7. When *a* increases and other variables are held equal, the *q*-intercept increases while the *p*-intercept stays the same: the ES curve is right-shifted toward higher real output. A similar proof applies to the converse case of an increase in *x* when other variables are controlled.

## Effort-Setting and Siphoning: a Single Transaction

Let the firm siphon capital-augmenting resources $\Delta k$ from the retail market so its production function becomes:

(A8) $\quad q = a \cdot \left(\bar{k} + \Delta k\right)^{1-\alpha} \cdot e^{\alpha}$ .

The ES curve with siphoning (the ES′ curve) becomes:

(A9) $$q = a \cdot \left(\bar{k} + \Delta k\right)^{1-\alpha} \cdot \left[1 - \lambda - \left(\frac{p - p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right]^{\alpha} .$$

**Proposition 4.** *The* ES′ *curve is concave to the origin: the marginal rate of substitution decreases as p increases.*

*Proof:*

Because $\Delta k$ is predetermined, its influence on equation A9 compared with equation A5 is purely scalar. Therefore, the proof suggested for proposition 1 is also applicable to this proposition. In other words, the signs of the first and second derivatives of $q$ with respect to $p$ for the ES′ curve are the same as those for the ES curve.

**Proposition 5.** *Siphoning shifts the* ES *curve toward real output.*

*Proof:*

Because $\Delta k$ is predetermined, its influence on equation A9 compared with equation A5 is identical to that of capital accumulation or a technological improvement, i.e. an increase in $a \cdot k^{1-\alpha}$ . Therefore, the proof suggested for proposition 3 is also applicable to this proposition, writing $a \cdot \left(\bar{k} + \Delta k\right)^{1-\alpha}$ for $a \cdot \bar{k}^{1-\alpha}$ in equation A7.

## Repeated Siphoning

The firm’s hidden revenues in repeated siphoning are given as:

(A10) $\quad b = r \cdot q \cdot (\hat{p} - p), \quad 1 > r > 0, \quad q > 0, \quad \hat{p} > p$

where

(A11) $\quad \hat{p} = z - \pi \cdot q, \quad z, \pi > 0$ .

The condition that $\hat{p} > p$ suggests:

(A12) $z > \pi \cdot q + p$ .

The first-order condition to maximize *b* with respect to *q* is:

(A13) $z = 2 \cdot \pi \cdot q + p$ .

From A12 and A13, positive hidden revenues ($b > 0$) should satisfy the condition:

(A14) $\pi \cdot q + p < z \leq 2 \cdot \pi \cdot q + p$ .

**Proposition 6.** *The* ES′ *curve has a negative slope: the partial derivative of q with respect to p is less than zero.*

*Proof:*

Combine equations A9, A10, and A11:

(A15) $$q = a \cdot \left[\bar{k}^{1-\alpha} + s \cdot r \cdot q \cdot \left(z - \pi \cdot q - p\right)\right]^{1-\alpha} \cdot \left[1 - \lambda - \left(\frac{p - p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right]^{\alpha}$$

Set $F = q - a \cdot \left[\bar{k}^{1-\alpha} + s \cdot r \cdot q \cdot \left(z - \pi \cdot q - p\right)\right]^{1-\alpha} \cdot \left[1 - \lambda - \left(\frac{p - p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right]^{\alpha}$ ,

$A = \left[1 - \lambda - \left(\frac{p - p_{-1}}{x \cdot p_{-1}}\right)^{\frac{1}{\beta}}\right]$, $B = \left(\frac{p - p_{-1}}{x \cdot p_{-1}}\right)$, and

$C = \left[\bar{k}^{1-\alpha} + s \cdot r \cdot q \cdot \left(z - \pi q - p\right)\right]$.

The implicit function theorem suggests:

$$\frac{\partial q}{\partial p} = -\frac{\frac{\partial F}{\partial p}}{\frac{\partial F}{\partial q}} .$$

$$\frac{\partial F}{\partial p} = a \cdot s \cdot r \cdot q \cdot \left(1 - \alpha\right) \cdot C^{-\alpha} \cdot A^{\alpha} + \frac{a \cdot \alpha}{x \cdot \beta \cdot p_{-1}} \cdot B^{\frac{1-\beta}{\beta}} \cdot C^{1-\alpha} \cdot A^{\alpha-1} > 0$$

and, because $z - 2 \cdot \pi \cdot q - p \leq 0$ according to A14,

$$\frac{\partial F}{\partial q} = 1 - a \cdot s \cdot r \cdot q \cdot (1 - \alpha) \cdot C^{-\alpha} \cdot A^{\alpha} \cdot \left(z - 2 \cdot \pi \cdot \cdot q - p\right) \geq 1$$

Therefore $\frac{\partial q}{\partial p} < 0$.

**Proposition 7.** *Under continuous siphoning the downward slope of the* ES′ *curve depends directly on that of the market demand curve.*

*Proof:*

This proposition is investigated through the effects on the *p*- and *q*-intercepts of the ES curve arising from an increase in the value of *π*, the downward slope of the market demand curve. The *p*-intercept is found by setting $q = 0$ in equation A15.

$$\text{(A16)} \quad p = p_{-1} \cdot \left[ x \cdot (1-\lambda)^{\beta} + 1 \right]$$

Equation A16 is equivalent to A6, suggesting hidden revenue does not change the *p*-intercept in repeated siphoning. In addition, the elasticity of market demand (the value of *π*) has no influence on the *p*-intercept.

The *q*-intercept is found from equation A15 by setting $p = p_{-1}$:

$$\text{(A17)} \quad q = a \cdot \left[ \bar{k}^{1-\alpha} + s \cdot r \cdot q \cdot (z - \pi \cdot q - p) \right]^{1-\alpha} \cdot (1-\lambda)^{\alpha}$$

Given positive hidden revenue ( $\pi \cdot q + p < z \le 2 \cdot \pi \cdot q + p$ ), other things being equal, when *π* decreases, the *q*-intercept increases: the ES curve becomes flatter because the *p*-intercept stays the same as before while the *q*- intercept increases. The same proof applies to the case when *π* increases.

**Proposition 8**. *Compared with the no-siphoning equilibrium, the continuous siphoning equilibrium will shift in favour of higher concealed inflation when the market demand curve is steeply sloping downward, and in favour of higher output in the converse case.*

*Proof:*

Equation A15 suggests that hidden revenues increase real output; however, real output is negatively associated with *π*, the slope of the market demand curve. When the market demand curve becomes steeper, real output decreases. Given the same nominal output target, i.e. the PS curve is the same, concealed inflation should therefore increase. The same logic applies when the market demand curve becomes flatter: the result is higher real output but less concealed inflation.

# Appendix 2. Demand Elasticities

*Table 2.1. The Price Elasticity of Market Demand Faced by the Firm in Postwar Soviet Industry by Branch*

| | Price elasticity of firm's market demand | Buyer can find other suppliers within the industry | Buyer can find near substitutes within the industry | Buyer can find near substitutes outside the industry |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Machinery | 0 | 0 | 0 | 0 |
| Light industry | 2 | 2 | 2 | 2 |
| Chemicals | 0 | 0 | 0 | 0 |
| Wood, pulp, and paper | 1 | 1 | 1 | 0 |
| Construction materials | 1 | 2 | 0 | 1 |
| Food industry | 2 | 2 | 1 | 2 |
| Ferrous metals | 0 | 0 | 0 | 1 |
| Electric power | 0 | 0 | 0 | 0 |
| Fuels | 1 | 0 | 1 | 1 |

*Sources and Notes.* 0 = "low" or "low probability," 1 = "medium" or "medium probability," 2 = "high" or "high probability." Column 1 is calculated as the median of figures in the three columns to the right. These take into account the structure of each industry (column 2), the scope for substitution among the industry's products (column 3), and the availability of alternative sources of supply (column 4); for more discussion see the text. Figures are based on our reading of a number of classic studies of the Soviet industrial and economic structure including Berliner (1952, 1957), Granick (1954, 1964), and Nove (1977), and of the Soviet "second economy" by Grossman (1977, 1979) and Katsenelinboigen (1977), as well as more recent studies such as Davies (1989), Davies (1994), Harrison (1996), Davis (1999). and Kim (2002). These studies support the rankings offered in the table rather than cardinal measures. In column 2 we index the ability of buyers to find alternative suppliers of a given product within the industry, setting it low in machinery, chemicals, ferrous metals, electricity, and fuels, and high in light industry, construction materials and the food industry, with other branches in between. In column 3 we index the ability of buyers to switch from the product they seek to another that is produced within the same industry, setting it low in machinery, chemicals, and construction materials, and high in light industry, with other branches in between. In column 4 we index the ability of buyers to switch from the product they seek to another that is produced in another industry or from another source such in-house sideline production or the second economy, setting it low in machinery, chemicals, wood, pulp, and paper, and electric power, and high in the light and food industries, with other branches in between.

# References


Ames, Edward (1965). *Soviet Economic Processes*, Homewood, IL: Irwin.

Belova, Eugenia (2001a). "Economic Crime and Punishment." In Paul R. Gregory, ed., *Behind the Façade of Stalin's Command Economy*, Stanford, CA: Hoover Institution Press, 131-58.

Belova, Eugenia (2001b). "Contract Enforcement Under Dictatorship: the Case of the Soviet Economy." PERSA Working Paper no. 14, University of Warwick, Department of Economics.

Belova, Eugenia, and Paul R. Gregory (2002). "Dictators, Loyal and Opportunistic Agents: the Soviet Archives on Creating the Soviet Economic System." *Public Choice*, 113(3-4), 265-286.

Berliner, Joseph S. (1952). "The Informal Organization of the Soviet Firm." *Quarterly Journal of Economics*, 66, 342-65.

Berliner, Joseph S. (1957). *Factory and Manager in the USSR*, Harvard University Press: Cambridge, MA.

Berliner, Joseph S. (1976). *The Innovation Decision in Soviet Industry*, Cambridge, MA: MIT Press.

Birman, Igor (1980). "The Financial Crisis in the USSR." *Soviet Studies*, 32, 84-105.

Central Intelligence Agency (CIA) (1979). *An Analysis of the Behavior of Soviet Machinery Prices, 1960-73: A Research Paper*, National Foreign Assessment Center, ER79-10631, Government Printing Office: Washington, DC.

Central Intelligence Agency (CIA) (1990). *Measures of Soviet Gross National Product in 1982 Prices*, Government Printing Office: Washington, DC.

Davies, R.W. (1989). "Economic and Social Policy in the USSR, 1917-41." In Peter Mathias and Sidney Pollard, eds, *The Cambridge Economic History of Europe*, vol. 8, *The Industrial Economies: The Development of Economic and Social Policies*, Cambridge University Press: Cambridge, 984-1047.

Davies, R.W. (1994). "Industry." In R.W. Davies, Mark Harrison, and S.G. Wheatcroft (editors), The Economic Transformation of the Soviet Union, 1913-1945, Cambridge University Press: Cambridge, 131-57.

Davies, R.W. (2001). "The Making of Economic Policy." In Paul R. Gregory, ed., *Behind the Façade of Stalin's Command Economy*, Stanford, CA: Hoover Institution Press, 61-80.

Davis, Christopher M. (1999). "Russia: a Comparative Systems Interpretation." In James Foreman-Peck and Giovanni Federico, eds, *European Industrial Policy: The Twentieth Century Experience*, Oxford University Press: Oxford, 319-97.

Dewatripont, Mathias, and Eric Maskin (1995). "Credit and Efficiency in Centralized and Decentralized Economies." *Review of Economic Studies*, 62(4), 541-55.

Ehrlich, Issac, and Francis Lui (1999). "Bureaucratic Corruption and Endogenous Economic Growth." *Journal of Political Economy*, 107(6), 270-93.

Ericson, Richard E. (1983). "On an Allocative Role of the Soviet Second Economy." In Padma Desai, ed., *Marxism, Central Planning, and the*

*Soviet Economy: Economic Essays in Honor of Alexander Erlich*, Cambridge, MA: MIT Press, 110-32.

Ericson, Richard E. (1984). “The ‘Second Economy’ and Resource Allocation under Central Planning.” *Journal of Comparative Economics*, 8(1), 1-24.

Goldfeld, Stephen M., and Richard E. Quandt (1988). “Budget Constraints, Bailouts and the Firm under Central Planning.” *Journal of Comparative Economics*, 12(4), 502-20.

Goldfeld, Stephen M., and Richard E. Quandt (1990). “Output Targets, Soft Budget Constraints and the Firm under Central Planning.” *Journal of Economic Behavior and Organization*, 14(2), 205-22.

Granick, David (1954). *Management of the Industrial Firm in the USSR*, Columbia University Press: New York.

Granick, David (1964). *Soviet Metal Fabricating and Economic Development*, University of Wisconsin Press, Madison, WI.

Granick, David (1987). *Job Rights in the Soviet Union: Their Consequences*, Cambridge: Cambridge University Press.

Gregory, Paul R. (1990). *Restructuring the Soviet Economic Bureaucracy*, Cambridge: Cambridge University Press.

Gregory, Paul R. (2001). “The Dictator’s Orders.” In Paul R. Gregory, ed., *Behind the Façade of Stalin’s Command Economy*, Stanford, CA: Hoover Institution Press, 11-33.

Gregory, Paul R., and Robert C. Stuart (2001). *Russian and Soviet Economic Performance and Structure*, 7th edn, Addison Wesley: Boston, MA.

Gregory, Paul R., and Aleksei Tikhonov (2000). “Central Planning and Unintended Consequences: Creating the Soviet Financial System, 1930-1939.” *Journal of Economic History*, 60(4), 1017-40.

Grossman, Gregory (1960). *Soviet Statistics of Physical Output of Industrial Commodities: their Compilation and Quality*, National Bureau of Economic Research and Princeton University Press, Princeton, N.J.

Grossman, Gregory (1977). “The ‘Second Economy’ of the USSR.” *Problems of Communism*, 26, 25-40.

Grossman, Gregory (1979). “Notes on the Illegal Private Economy and Corruption.” In U.S. Congress Joint Economic Committee, *Soviet Economy in a Time of Change*, vol. 1, U.S. Government Printing Office: Washington DC, 834-55.

Grossman, Gregory (1986). “Inflationary, Political, and Social Implications of the Current Economic Slowdown.” In Hans-Hermann Höhmann, Alec Nove, and Heinrich Vogel, eds, *Economics and Politics in the USSR: Problems of Interdependence*, Boulder, CO: Westview Press, 172-97.

Grossman, Gregory (1998). “Subverted Sovereignty: Historic Role of the Soviet Underground.” In Stephen S. Cohen, Andrew Schwartz, and John Zysman, eds, *The Tunnel at the End of the Light: Privatization, Business Networks, and Economic Transformation in Russia*, Berkeley, CA: University of California, 24-50.

Harrison, Mark (1996). *Accounting for War: Soviet Production, Employment, and the Defence Burden, 1940-1945*, Cambridge: Cambridge University Press.

Harrison, Mark (1998). “Prices, Planners, and Producers: an Agency Problem in Soviet Industry, 1928-1950.” *Journal of Economic History*, 58(4), 1032-62.

Harrison, Mark (2000). “Soviet Industrial Production, 1928 to 1955: Real Growth and Hidden Inflation.” *Journal of Comparative Economics*, 28(1), 135-55.

Harrison, Mark (2001). “Providing for Defense.” In Paul R. Gregory, ed., *Behind the Façade of Stalin’s Command Economy*, Stanford, CA: Hoover Institution Press, 81-110.

Harrison, Mark (2002). “Coercion, Compliance, and the Collapse of the Soviet Command Economy.” *Economic History Review*, 50(3), 397-433.

Harrison, Mark, and Nikolai Simonov (2000). “Voenpriemka: Prices, Costs, and Quality in Defence Industry.” In Mark Harrison and John Barber, eds. *The Soviet Defence Industry Complex from Stalin to Khrushchev*, London and Basingstoke: Macmillan, 223-45.

Hodgman, Donald R. (1954). *Soviet Industrial Production, 1928-1951*. Harvard University Press: Cambridge, MA.

Holzman, Franklyn D. (1960). “Soviet Inflationary Pressures, 1928-1957: Causes and Cures.” *Quarterly Journal of Economics*, 74(2), 167-188.

Kaplan, Norman M., and Richard Moorsteen (1960). “Indexes of Soviet Industrial Output.” 2 vols., Research Memorandum RM-2495, The RAND Corporation,: Santa Monica, CA.

Kaser, Michael (1975). “The Economy: A General Assessment.” In Archie Brown and Michael Kaser, eds, *The Soviet Union since the Fall of Khrushchev*, London: Macmillan, 196-217.

Katsenelinboigen, Aron (1977). “Coloured Markets in the Soviet Union.” *Soviet Studies*, 29(1), 62-85.

Khanin, Grigorii I. (1991). *Dinamika ekonomicheskogo razvitiia SSSR*, Nauka: Novosibirsk.

Kim, Byung-Yeon (1997). “Soviet Household Saving Function.” *Economics of Planning*, 30(2-3), 181-203.

Kim, Byung-Yeon (2002). “Causes of Repressed Inflation in the Soviet Consumer Market: Retail Price Subsidies, the Siphoning Effect, and the Budget Deficit.” *Economic History Review*, 55(1), 105-127.

Kornai, János (1980), *The Economics of Shortage*, 2 vols, Amsterdam: North-Holland.

Kornai, János (1992), *The Socialist System: The Political Economy of Communism*, Clarendon: Oxford.

Kroll, Heidi (1986), “Decentralization and Precontract Disputes in Soviet Industry.” *Soviet Economy*, 2(1), 51-71.

Kudrov, Valentin M. (1997). *Sovetskaia ekonomika v retrospektive: opyt pereosmyslenii*, Nauka: Moscow.

Lushin, S. (1990). “Material’no-finansovaia sbalansirovannost’ narodnogo khozyaistvo.” *Finansy SSSR*, no 5, 24-31.

Mauro, Paolo (1995). “Corruption and Growth.” *Quarterly Journal of Economics*, 110(3), 681-712.

Moorsteen, Richard (1962). *Prices and Production of Machinery in the Soviet Union, 1928-1958*, Harvard University Press: Cambridge, MA.

Nove, Alec (1958). “The Problem of ‘Success Indicators’ in Soviet Industry.” *Economica*, 25(97), 1-13.

Nutter, G. Warren (1962). *The Growth of Industrial Production in the Soviet Union*. National Bureau of Economic Research and Princeton University Press: Princeton, NJ.

Qian, Yingyi (1994). "A Theory of Shortage in Socialist Economies Based on the 'Soft Budget Constraint'." *American Economic Review*, 84(1), 145-56.

Sah, Raaj Kumar (1991). "Fallibility in Human Organizations and Political Systems." *Journal of Economic Perspectives*, 5(2), 67-88.

Schroeder, Gertrude E. (1985). "The Slowdown in Soviet Industry, 1976-1982." *Soviet Economy*, 1(1), 42-74.

Shleifer, Andrei, and Robert W. Vishny (1992). "Pervasive Shortages Under Socialism." *RAND Journal of Economics*, 23(2), 237-46.

Shleifer, Andrei, and Robert W. Vishny (1993). "Corruption." *Quarterly Journal of Economics*, 108(3), 599-617.

Treml, Vladimir G., and Michael Alexeev (1994). "The Growth of the Second Economy in the Soviet Union and its Impact on the System." In Robert W. Campbell, ed., *The Postcommunist Economic Transformation*, Boulder, CO: Westview Press, 221-247.

Woodruff, David (1999). *Money Unmade: Barter and the Fate of Russian Capitalism*, Ithaca, NY: Cornell University Press.