IMF Staff Papers
Vol. 48, Special Issue

Falling Tax Compliance and the Rise of the Virtual Budget in Russia

BRIAN AITKEN*

The decline in cash revenue in Russia has been the key macroeconomic policy failure of the transition. During 1994–98, a sharp deterioration in cash compliance was offset by a rise in non-cash revenue, as the government increasingly financed its spending through mutual arrears write-offs. This paper argues that the fall in cash compliance emerged when money printing was replaced with a method of budget financing that did not, in the short run, compromise the government's goals of low inflation, a stable exchange rate, and low interest rates, but which ultimately has led the government into a low cash revenue trap. [JEL: H26, H30, E62, E65, O23]

Russia's macroeconomic policy has been dogged for years by the federal government's failure to collect enough cash revenue. Cash collection has fallen short of budget targets as an annual event, and every year the large budget gaps have been filled by spending arrears and additional short-term borrowing. The combination of growing debt and shrinking revenues naturally alarmed creditors, and it is not surprising that confidence ultimately gave way in the summer of 1998 with disastrous consequences for the government.

What is surprising though, given what was at stake, is that the government did not act more forcefully to reverse the revenue decline. The government clearly understood that its major achievements—low inflation and a predictable exchange rate—were increasingly jeopardized by chronically low revenue. Why then was it

*Brian Aitken is a Senior Economist at the International Monetary Fund. The author thanks John Anderson, Tom Richardson, Tapio Saavalainen, Ratna Sahay, and Jeromin Zettelmeyer for comments on a previous draft, and participants of the December 1998 European II Department seminar.

seemingly impotent to reverse the revenue decline? Did technical barriers keep revenue low, or was collecting taxes merely a question of "political will"? If the reasons were not technical, why would the government have allowed political considerations to dictate a policy that was clearly selfdestructive?

In this paper, I argue that the revenue decline was brought about by a budget practice that was actually quite sensible from a short-run perspective—namely, the virtual budget, or the practice of offsetting tax and expenditure arrears. The rise of the virtual budget, however, triggered a fundamental shift in taxpayer compliance as enterprises discovered that they could benefit from tax offsets by accumulating tax arrears. Falling cash compliance further aggravated the government's need for financing, which, ironically, made offsets appear even more attractive as a short-run solution. This set in motion a vicious circle where the expectation that tax arrears would be offset became a self-fulfilling prophesy. With the practice of offsets firmly entrenched, expectations could be changed, and compliance improved, only through a long and painful campaign against tax delinquents with a great short-run loss in non-cash budget financing. In this sense, the government was caught in a low cash revenue trap. By mid-1998, it was simply too late for fresh revenue raising efforts to prevent devaluation and default.

The key event setting this process in motion was the elimination of the inflation tax, when the government stopped printing money but did not support this with adequate cuts in expenditure. Instead of more drastic expenditure cuts, the government sought to replace money printing with budget financing that would not in the short run compromise its goals of low inflation, a stable exchange rate, and low interest rates.

The revenue decline is described in several stages. First, a fundamental change in taxpayer behavior took place beginning in 1995 as cash compliance, which had been more or less constant during 1992–94, fell sharply. Second, the elimination of the inflation tax led to the rise in the virtual budget. I argue that the virtual budget reduced the incentives for enterprises to pay taxes as well as for the government to punish tax delinquents. Finally, a simple model of tax compliance is developed to show how these developments can be characterized as a low revenue trap.

In considering the arguments in this paper, it is useful to keep in mind at least one possible alternative explanation for the revenue decline based loosely on the Gaddy and Ickes (1998) model of the virtual economy.[1] In this alternative, enterprises don't choose cash compliance, except perhaps at the margin. Instead, they can be thought to pay in cash only "what they can." If during high inflation, for example, seigniorage was transferred from households to enterprises in the form of budget subsidies, then eliminating the inflation tax might have reduced the pool of real resources from which enterprises could pay taxes. Cash revenue might be expected to fall even if the measured tax liability, which is a reflection of virtual economic activity, did not. While I do not reject this alternative, I discuss some reasons why it is not an entirely persuasive explanation of the revenue decline. In addition, I discuss what implications it might have for reversing the decline.

[1]Gaddy and Ickes (1998). In their paper, Gaddy and Ickes do not model tax compliance directly, but argue that cash revenue collection can be increased only by redirecting enterprises' cash proceeds away from other expenses such as wages.

I. Federal Revenue Developments

In this section, I focus on tax revenue developments of the federal government. For data reasons, I set aside questions regarding local government revenues (comprised of 89 separate regions) and payroll tax revenues accruing to the pension fund, social insurance fund, and other extrabudgetary funds.[2]

Federal cash revenue has declined sharply since 1992 (Table 1). Much of the decline during the period 1992–94 is explained by explicit changes in tax policy, such as a reduction in the VAT and profits tax rates and a phasing out of export duties. These changes are also reflected in a decline in the assessed tax liability, which can be estimated quite simply using data on tax arrears as assessed by the state tax service and confirmed by enterprises, including deferrals but excluding fines and interest penalties.[3] Since tax policy changes during 1992–94 are discussed elsewhere, I will not focus on them in this paper.[4]

After 1994, however, the revenue decline took on a very different character. During this period, cash collections fell by as much as 2½ percent of GDP despite a modest *increase* in the assessed tax liability (see also Figure 1); the share of the assessed liability paid in cash—or cash compliance—fell from 89 percent in 1994 to 65 percent in 1996. Despite this fall, total revenue including non-cash receipts has been steady since 1994. The figures for non-cash revenue simply reflect the quantity of tax arrears cleared through one of the government-sponsored offset schemes, which I will discuss in more detail below.

II. Implications for Common Explanations of the Poor Revenue Performance

Before turning to the reasons for the deterioration in cash compliance and the rise in non-cash revenue, it is important to note what the stability of the tax liability implies about many of the commonly cited explanations for the poor

[2]Data on local government and extrabudgetary revenues are especially difficult to interpret. Local governments are widely believed to systematically underreport their revenue, mainly because the distribution of federal transfers is in part determined by local revenue shortfalls. In addition, an unknown share of revenues and expenditures reported by local governments and some extrabudgetary funds take place through non-cash transactions such as through mutual arrears clearing operations. While the share of these operations is believed to be large and growing, the absence of a time series prevents cash and non-cash revenue developments from being analyzed separately. Aside from these concerns, however, there is no fundamental reason why the arguments in this paper cannot be extended to local and extrabudgetary revenues.

[3]Total revenue is based on official Russian treasury figures excluding privatization proceeds and proceeds from sales of stocks of precious metals, and is composed of cash revenue and non-cash revenue (tax offsets). Tax offsets are recorded as revenue in the treasury figures when the offset transaction takes place, rather than when the tax liability was incurred. Data on changes in the stock of tax arrears come from the state tax service, and exclude fines and penalties. The total tax liability as assessed by the state tax service is estimated by adding total revenue to the change in the stock of tax arrears. An "adjusted tax liability," shown in Table 1, is estimated by adjusting for differences in timing between when a taxable transaction actually took place and when a tax liability was later formally incurred as part of an offset operation.

[4]See for example, Chapter 5 of Citrin and Lahiri (1995).

Table 1. Russia: Federal Revenue Developments, 1992–98

(Percent of GDP)

	1992	1993	1994	1995	1996	1997	1998	
							Jan.–Jun.	Annualized[1]
Assessed tax liability (A+B)	16.6	15.3	12.9	13.4	13.6	13.9	13.5	13.4
A. Revenue	16.6	13.7	11.8	12.2	11.9	12.6	11.4	10.6
Cash	16.6	13.7	11.4	10.8	8.9	9.4	9.9	9.9
Non-cash (tax offsets)[2]	0.0	0.0	0.4	1.4	3.0	3.1	1.4	0.7
B. Change in stock of tax arrears[3]	. . .	1.6	1.1	1.3	1.7	1.3	2.1	2.8
Memorandum items:								
Cash as a share of assessed liability (in percent)	. . .	90	89	80	65	68	74	74
Adjusted tax liability[4]	16.6	15.6	13.4	13.9	13.7	13.8	...	. . .
Stock of federal tax arrears	0.0	1.6	1.5	1.8	3.0	3.9	4.8	. . .
GDP (billions of rubles)	19.2	172	611	1630	2256	2675	1242	2700

Sources: Russian Treasury, state tax service, and author's calculations.

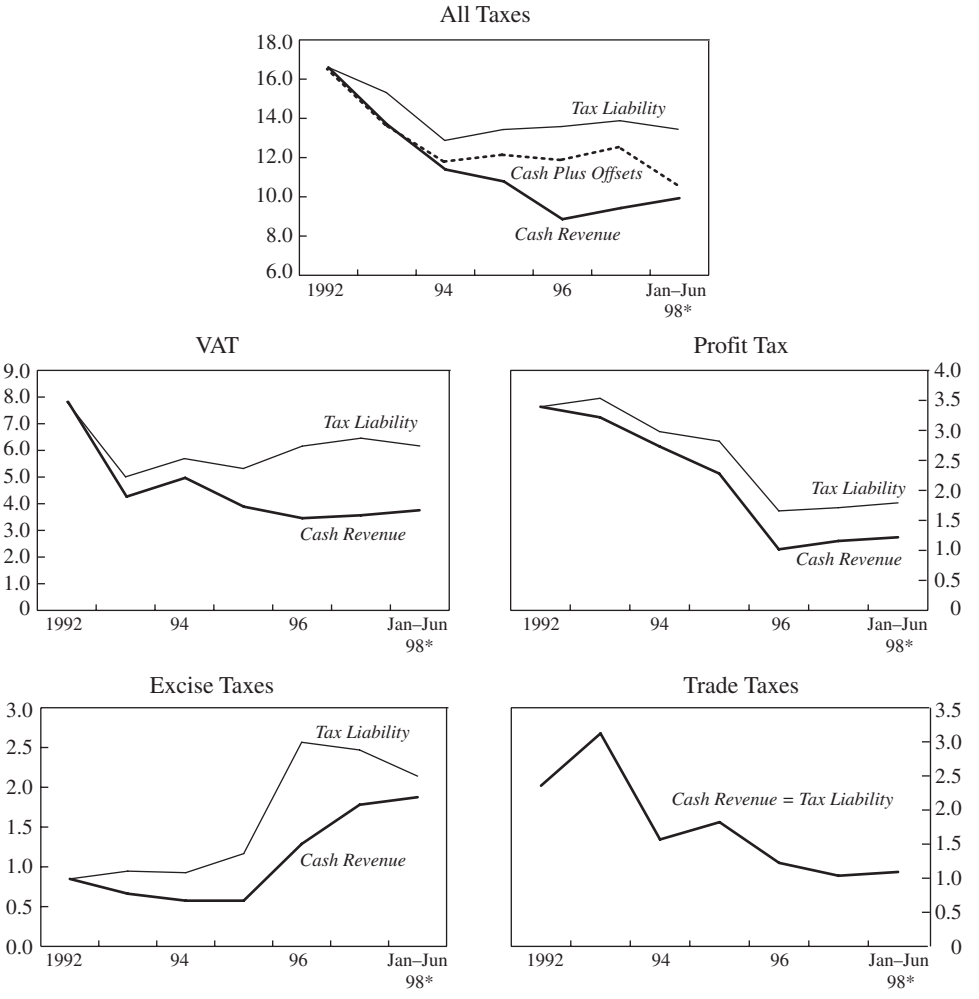
[1]Based on developments, seasonally adjusted, through June 1998.

[2]Tax offsets are reflected in full when the operation is conducted, rather than according to the budget year in which they are booked. Therefore, total revenue will differ from IMF program definitions and from Russian Ministry of Finance definitions.

[3]Excluding fines and penalties.

[4]Adjusted to correct for the bias in measured tax liabilities caused by new tax liabilities arising from offset operations.

Figure 1. Assessed Federal Tax Liability and Cash Revenue
(Percent of GDP)



Sources: Russian Treasury, state tax service, and author's calculations.

revenue performance in Russia. For this, it is first useful to decompose cash revenue in the following way.

$$\begin{array}{lllllllll} \textit{cash} & \textit{tax} & & \textit{cash} & & \textit{tax} & & \textit{observed} & \textit{cash} \\ \textit{revenue} = & \textit{liability} & \times & \textit{compliance} & = & \textit{rate} & \times & \textit{tax base} & \times \ \textit{compliance} \end{array}$$

Here cash compliance is taken to mean, by definition, the share of the tax liability as assessed by the tax service that is paid in cash. Changes in cash revenue will reflect changes in tax rates, the observed base (as measured by the tax service), or cash compliance. Anything that affects tax rates or the observed base

will also affect the assessed tax liability. For example, a reduction in the value-added tax (VAT) rate, an increase in VAT exemptions, or an increase in under-reporting of economic activity will all lead to a decline in the assessed VAT liability. If we see that the assessed VAT liability has been stable, then none of these factors could cause a decline in overall cash revenue.

In fact, there is a long list of commonly cited reasons for the poor revenue performance in Russia that fail to explain the recent revenue decline. Many of the reasons focus on a shrinking of the measured tax base, either due to exemptions and tax concessions or through tax evasion related to underreporting of production and the growth of the informal economy.[5] But, in order for these factors to explain the revenue decline after 1994, they would require a corresponding fall in the measured tax liability.

Likewise, changes in tax policy also fail to explain the revenue decline after 1994. As is clear from Figure 1, for example, the liability for some taxes, such as export taxes and the profits tax, declined after 1994. These changes, however, were offset by an increase in the liability for other taxes such as excises and VAT. More important for the decline in overall cash revenue, cash compliance for each of the major taxes worsened beginning in 1995.

Finally, widely recognized tax administration problems such as the complexity of the tax system and a dysfunctional state tax service could have played a role in the revenue decline only insofar as they explain why assessed taxes were not collected. A corrupt, poorly focused, poorly trained, and poorly funded tax service, an administratively complex VAT and profits tax, a lack of established audit procedures, and a proliferation of transfer pricing might have hindered correct tax assessment and contributed to tax evasion, but revenue developments show that problems assessing tax liabilities cannot be driving the revenue decline.

To summarize, there are many factors that might be used to argue why revenue was not higher than it in fact was, or why Russia has failed to achieve its full revenue potential. But the stable tax liability implies that the fall in taxpayer compliance must be the driving force behind the revenue decline after 1994.

III. The Rise in the Virtual Budget

One of the notable features of the decline in cash compliance is that it was more or less offset by a rise in non-cash revenue, but this does not mean that the government simply began allowing enterprises the choice of paying taxes in cash or in kind. In fact, I argue that the decline in cash compliance was caused by the government's desire to maintain expenditure following the elimination of the inflation tax. I make this argument in two stages. First, in this section, I show that the rise in non-cash revenue, and correspondingly non-cash expenditure, emerged when the government substituted one form of soft financing—-money printing—for another—tax offsets. Second, I argue that the virtual budget, or the practice of

[5]See, for example, OECD (1997), Ickes, Murrell, and Ryterman (1997).

carrying out expenditures and collecting revenues in non-cash form, created an environment in which enterprises benefited greatly by reducing their cash compliance.

Throughout the period 1992–97, cash spending fell well short of the budgeted level (Table 2). Since spending commitments were made based on the budgeted limits, the shortfall in cash spending led directly to an increase in federal government spending arrears. Prior to 1995, these arrears were mostly cleared in cash within the spending limits of the following year. At that time, clearing arrears was not especially difficult since (1) high inflation significantly reduced the cost to the budget in real terms of clearing arrears in the following year; and (2) spending in excess of identified financing was generously covered by inflationary money creation. By the end of 1993, for example, the federal government had run Rub 4 billion in spending arrears, or about 2.3 percent of 1993 GDP, which were to be cleared within the spending limits of the 1994 budget. These arrears amounted to less than ⅔ of 1 percent of GDP in 1994. High inflationary financing allowed the government to run arrears each year and at the same time keep the real stock of arrears from growing.

Budget practices shifted in 1995, when the government strengthened its efforts to control inflation by sharply reducing monetary financing of the budget (Table 2).

Table 2. Russia: Shortfall in Federal Government Expenditure Relative to Budget, 1992–97
(Percent of GDP)

	1992	1993	1994	1995	1996	1997
Non-interest expenditure						
Budget	. . .	20.6	22.7	14.9	16.3	16.5
Cash execution	26.0	18.3	20.8	12.6	11.2	11.7
Cash shortfall relative to budget	. . .	2.3	1.9	2.3	5.1	4.8
Memorandum items:						
Monetary financing of the federal budget	. . .	6.5	9.7	2.7	2.3	1.3
Annual average inflation (in percent)	1735	875	307	197	48	15

Sources: Russian Treasury, state tax service, and author's calculations.

Keeping inflation down meant that spending arrears could no longer be cleared by printing money. To break with this practice the federal government chose to clear its stock of 1994 arrears by issuing special treasury obligations (KOs) to suppliers. The cash burden of repaying these obligations was minimized by giving suppliers the option of using them to clear outstanding tax arrears, which when cleared were booked as non-cash revenue.

As a one-time measure, issuing special treasury obligations was a convenient way of clearing past arrears. However, spending arrears continued to mount in 1995; even though budgeted expenditures were reduced sharply in that year, cash spending fell short of the budgeted level by nearly 2½ percent of GDP. In addition

to the measured shortfall, unbudgeted spending related to the war in Chechnya, the run-up to parliamentary elections in late 1995, and presidential elections in mid-1996 added to the cash shortfall.

As spending arrears mounted, the federal government came under heavy pressure to "clear the books" once again by offsetting these arrears against the tax arrears of enterprises. The result was a succession of mutual arrears clearance schemes, or tax offsets. The mechanics of each offset scheme differed, but in all cases the federal budget cleared its arrears to suppliers by issuing a claim on the government that could be used to pay tax arrears.[6] Each scheme was intended only as a bookkeeping exercise to clear past—and not current—arrears, but the sequential and ongoing nature of the schemes meant that they de facto financed current expenditures (Table 3). By 1996, over one-fourth of total federal revenue took the form of non-cash tax offsets.

Table 3. Russia: Federal Government Expenditures Financed Through Offsets, 1992–97
(Percent of GDP)

	1992	1993	1994	1995	1996	1997
Expenditures financed through offsets						
by budget year	0.0	0.0	1.6	3.2	3.9	2.2
by date of offset transaction	0.0	0.0	0.4	1.6	3.0	3.1
Memorandum Item:						
Total federal non-interest expenditures	26.3	18.3	21.2	14.6	14.6	14.4

Sources: Russian Treasury, state tax service, and author's calculations.

IV. The Virtual Budget and the Fall in Cash Compliance

The emergence of the virtual budget triggered a fall in cash compliance for two reasons: (1) tax offsets increased the benefits to enterprises of running tax arrears; and (2) tax offsets reduced the government's incentive to penalize tax delinquents.

Impact of Offsets on Taxpayer Incentives

As I have noted, each of the offset schemes was intended to clear only past tax arrears and did not relieve taxpayers of the need to pay current tax liabilities in cash. Repeated episodes of "one-time," "final" offset schemes confirmed enterprises' expectation, however, that, as long as the budget was running arrears, there would be another offset scheme in the future. To participate in a future offset, an enterprise needed tax arrears.

[6]See Crotty and others (1997) for a complete description of the various offset schemes.

It is clear why enterprises supplying state orders might prefer to settle their tax liabilities effectively "in kind" through mutual arrears clearing rather than in cash. Many of the enterprises supplying the government, such as in the defense sector, have few alternative markets for their goods. The prices these enterprises charge the government are commonly believed to be well in excess of the cash value of the goods in a competitive market. Selling goods to the government on credit, and then using this credit to write off tax liabilities, was one way for enterprises to realize profits on their less marketable goods. Even profitable enterprises, such as Gazprom, also benefited from trade in kind with the government. Pipeline constraints effectively limited the supply of natural gas to the export market where Gazprom's profits were greatest and customers paid in cash. In contrast, the cash value of supplying an additional unit of gas to the domestic market was quite low, far below the contract price that Gazprom was able to charge for supplying gas to government agencies. With low marginal costs of production, Gazprom had a clear incentive to build up tax arrears on its exports, and then clear those arrears by offsetting them against its sales to the government.

While it is not surprising that enterprises that were owed money by the budget would run up tax arrears, a crucial feature of the various offset schemes is that they extended the potential benefits of tax arrears to *all enterprises* whether or not enterprises had any direct link to the federal budget. In all the offset schemes, the enterprise holding the claim on the budget could, whether or not it was explicitly allowed, trade this claim to another enterprise wishing to use the claim to clear tax arrears. In some schemes the claims were securitized, and an active and highly lucrative secondary market quickly developed. Other schemes required chains of debtors and creditors, with the budget on one end and the tax delinquent on the other. In these cases, claims were effectively traded through bogus post-dated contracts and side agreements between firms.

The main point is that offsets reduced the incentive to pay taxes in cash for all enterprises in the economy. Offsets were particularly valuable to the more politically exposed large enterprises. The Karpov Commission found that in their sample of the 210 largest tax delinquents, less than 8 percent of payments to the federal budget during 1996 and the first half of 1997 took the form of cash.[7] Likewise, the potential gains from offsets were particularly large for enterprises with profitable cash markets but no direct claim on the budget. As such, it is no surprise that the tax arrears of the oil sector rose sharply in advance of the widely anticipated offset exercise that took place in late 1997 (Table 4).

The Government's Response to Falling Tax Compliance

One might expect that the government, having lost the inflation tax as a source of financing, would do everything in its power to improve tax compliance. After all, the government knew precisely which enterprises owed what taxes, and the list of the largest tax delinquents was publicly known. Why then did the government tolerate the *decline* in compliance? This is somewhat of a paradox and critical in

[7]Report of the Interdepartmental Balance Commission (1997).

Table 4. Russian Federation: Federal Tax Arrears of the 100 Largest Tax Debtors, by Sector

(End of period stock in billions of rubles)

	December 1995	July 1996	October 1996	April 1997	Sepember 1997	December 1997	April 1998	September 1998
Auto	1.8	4.0	4.4	4.7	5.2	5.9	5.7	5.3
Coal	0.1	0.1	—	0.2	0.6	0.6	1.0	0.7
Energy	2.3	2.6	2.3	4.0	5.7	5.7	5.0	6.7
Gas	1.9	6.2	6.8	7.9	5.1	8.0	10.5	16.7
Oil	5.6	10.4	12.3	10.5	13.1	9.1	4.6	5.3
Rail	2.6	2.4	2.6	2.2	3.6	3.9	3.3	4.9
Other	0.4	0.2	0.4	1.0	0.3	0.3	0.7	3.8
Total	15.6	27.2	29.6	32.7	35.5	35.6	33.5	43.2
Memorandum item:								
Total federal tax arrears	30	58	74	82	96	104	120	146

Sources: Russian Treasury, state tax service, and author's calculations.

explaining why cash compliance actually fell. The answer is that in the short run the government found tax offsets to be a less costly alternative, from both the political and financial points of view, to taking steps against tax delinquents. Enforcing cash compliance is politically costly to the government. Sanctions against tax delinquents such as seizing assets and bankrupting enterprises can be effective only if they inflict pain on owners and managers, either by stripping ownership or shutting down the operations of the enterprise. By their nature then, sanctions will involve some short-run political cost to the government; owners and managers are often politically powerful, particularly when closely allied with regional governments, and displaced workers are not likely to be sympathetic to the federal government's cause. But if threats against tax delinquents are to be credible, taxpayers need to perceive the government as willing to carry out these threats to their conclusion despite the political costs.

At the same time, the sanctions against taxpayers rarely provide any direct short-run financial benefit to the government. One potentially cost-effective sanction would be seizing the liquid assets of delinquent enterprises. However, enterprises have increasingly resorted to barter and use of money surrogates, such as veksels and arrears, and as a result rarely have any liquid assets to seize.[8] In the Karpov Commission's sample of large tax delinquents, for example, some 57 percent of the enterprises received less than 20 percent of their total revenue in cash form. Based on this sample, the Commission concluded that ". . . it makes no sense to apply fiscal punishment measures directed at finding and arresting accounts of the majority of 'classical' industrial enterprises. Indeed, there is

[8]The state tax service has the legal authority to block a tax delinquent's bank accounts and force all current revenues into a single settlement account that is applied to the payment of tax debts. Hendley, Ickes, and Ryterman (1998) discuss the increasing use of barter and other non-cash transactions as a way of evading seizure of revenues through the single settlement account system.

nothing there to find."[9] I set aside the question of why these enterprises are cash starved, and there certainly is evidence to suggest that they are cash starved *by design*. In any case, seizing liquid assets of tax delinquents is rarely an effective option.

As an alternative, the tax authorities on occasion seize non-liquid productive assets of enterprises. For legal and political reasons, however, an enterprise's assets can only be seized piece by piece, rather than as a coherent whole.[10] Moreover, until the federal debt center was established in mid-1998 there was not even a centralized market outlet for the sale of seized property. For these reasons, seized assets have had very little cash value. Indeed, seizing physical assets has been such a logistical headache that the tax authorities have been known to declare assets "seized" without bothering to remove them from enterprise premises. The result has been that, regardless of the book value of the assets seized, the cash proceeds from seizure and sale of productive assets have been a negligible fraction of the volume of tax arrears cleared through offsets.

Tax offsets, on the other hand, offered an alternative to sanctions that made much more sense to the government from a short-run perspective. Offsets were a way of capturing the entire nominal value of outstanding tax arrears and channeling it toward the clearance of the already large stock of expenditure arrears of the budget. At the same time, offsets avoided the political costs associated with sanctions. From a narrow, short-run point of view, they appeared to all parties to be a rational solution to a complex set of problems. Enterprises understood this and, as a result, did not take seriously government threats that painful sanctions would be taken against tax delinquents.

This explains why a number of extraordinary steps taken by the government to improve taxpayer discipline generally failed. In late 1996 the government established the Emergency Tax Commission (VChK) to review in a high profile manner the tax liabilities of some of the largest tax debtors. Only in a few cases, however, did the commission take actions to effectively penalize the enterprises for failing to clear tax debts. In most cases, consequences such as asset seizure and bankruptcy were avoided through agreements whereby the enterprises would stay current on new tax liabilities and restructure old tax debts. When declarations of tough action were made by the VChK, they were rarely carried out to their conclusion. In fact, the high profile nature of the VChK's failure to follow through on its threats could well have contributed further to the decline in cash compliance (Crotty, and others, 1997).

Several other high profile efforts also proved ineffective in the face of strong political opposition. In mid-1998, the tax service launched a public campaign to seize assets of Gazprom subsidiaries, which had been accumulating tax arrears at

[9]Interdepartmental Balance Commission (1997).

[10]In December 1997, the Emergency Tax Commission took a path-breaking decision to seize assets of several major tax delinquents in the form of seizure of their entire "property complex." This decision, by preserving an enterprise's value-producing operation in tact, would have greatly increased the benefits to the government of seizure as well as increasing the costs to the enterprise managers and owners. For these reasons, the decision mobilized the direct opposition of powerful political lobby groups and was very soon reversed on legal grounds.

a rapid pace since the end of 1997 (see Table 4). However, the campaign mobilized widespread political opposition against the government, and in the end an agreement was reached whereby Gazprom would pay only a part of its current tax liability in cash. As a result, Gazprom's tax arrears continued to mount in the second half of 1998. Likewise, an attempt by the government in mid-1998 to restrict export access for oil companies with growing tax arrears ran directly against the interests of powerful lobby groups, and the measure was largely circumvented in the end.

The failure to enforce tax collection is also reflected in the daily deliberations regarding lower profile tax delinquents. Of over 1,400 tax delinquent enterprises reviewed by the regional Interbranch Balancing Commissions during the first six months of 1998, in only 5 percent of the cases was a decision taken to seize assets. Although bankruptcy was initiated in about 19 percent of the cases, the poor state of the bankruptcy law and the quality of the courts prevent bankruptcy from being regarded as a serious threat. In contrast, in about 40 percent of the cases the tax debts were either deferred or restructured, or the balancing commission postponed a decision.[11]

Reasons for the Fall in Cash Compliance

To summarize, the fall in cash compliance was the product of (1) expenditure arrears of the budget and (2) a shift in the method of financing these arrears from money printing to tax offsets. The method of financing contributed to the fall in compliance by rewarding tax delinquents, while the government's short-run need to finance expenditure arrears assured tax delinquents that their behavior would be tolerated. Since both factors contributed to the decline in cash compliance, simply declaring an end to tax offsets should not by itself lead to a recovery in cash revenue. So long as the government is spending in arrears, enterprises are likely to continue running tax arrears on the expectation that some form of mutual arrears clearing exercise will need to take place.

V. The Low Revenue Trap in a Model of Cash Compliance

In this section, I bring together the factors that I have argued led to the fall of cash compliance into a coherent framework. I describe a simple game between the government—by which I am referring to the fiscal authorities—and enterprises in which cash compliance is induced by the threat that the government will seize assets (see Appendix for a complete description). In deciding how to respond to nonpayment the government determines a seizure policy balancing the cash revenue gained from seizing assets against the alternative of using unpaid tax liabilities as an imperfect substitute for cash revenue as a means of financing expenditure needs.

This model serves two modest but useful purposes. First, it describes clearly in a static framework the main argument of the paper—that a loss of cash

[11]Report of the Interdepartmental Balancing Commission (1998).

financing paradoxically increases the government's willingness to tolerate lower cash compliance. The second purpose is to illustrate how, over time, the government might find itself caught in a low cash revenue trap. In other words, the government's short-run goal of financing its expenditures *from any means* might outweigh the longer-run desire to finance expenditures *from current cash revenue*, leading it to postpone the costly investment of establishing a reputation for not tolerating tax delinquency. If enterprises understand this, they will challenge the government's commitment to enforce cash collection and reduce their compliance, which will in turn further aggravate the government's short-run financing concerns. While the model is fairly simple, it nonetheless provides a basic framework for a richer model to examine these dynamic aspects more thoroughly.

The Government's Preferences

The model begins with the assumption that the government (again, taken to mean the fiscal authorities) wishes to finance as much as possible of a given target expenditure level. In making this assumption, the larger question of how this level is determined in the first place is deliberately overlooked. This is an important question, and no attempt is made to answer it here. For the moment, one can assume that the fiscal authorities are essentially given the expenditure target, which itself is determined by larger political imperatives.

Assume expenditures can be financed from three sources: (1) cash revenue collected from enterprises; (2) taxes collected in non-cash form through offsets; and (3) other sources such as seigniorage or borrowing, which are taken to be exogenous. By treating them as exogenous, I am assuming that these sources are determined by factors outside the model, such as a given inflation target in the case of seigniorage or, in the case of borrowing, the desire to keep domestic real interest rates low and limit the growth of government debt. As with the target expenditure level, I do not intend to argue that these sources of financing are outside the control of the government but are simply limited by other policy goals, which are taken to be given for the purposes of the model. The model will then allow us to consider the effects of changes in other sources of financing, for a given target expenditure level, on cash and non-cash revenue collection.

A key aspect of the model's setup is the value the government places on cash revenue compared with other sources of financing. All else equal, the government would prefer as much financing as possible to come from cash revenue. From a long-run perspective the solvency of the government depends on its ability to finance expenditures with revenue, as opposed to borrowing or unsustainable money printing. With regard to revenue itself, the government has a clear preference for cash revenue over non-cash revenue. One reason is that some expenditures, such as wage payments and debt service, simply cannot be paid in kind or through offsets. Another is that cash is more fungible and does not distort government expenditure priorities toward spending on goods that tax delinquents happen to supply.

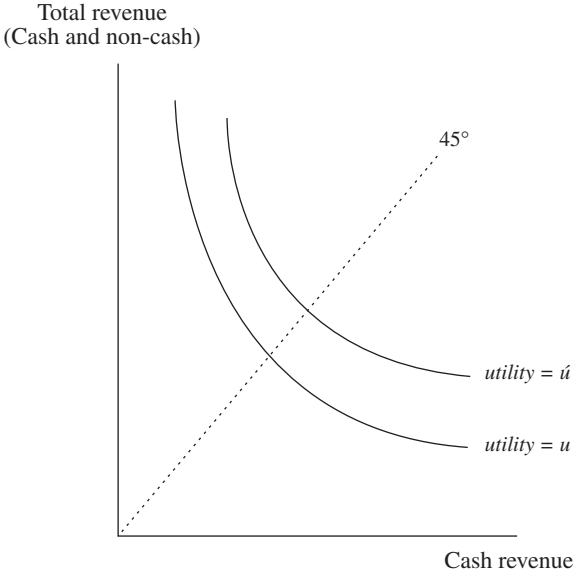
While the government prefers cash revenue, it nonetheless has a desire to finance its target expenditure level, regardless of the source of financing. In other

words, the government would not be happy with a relatively high level of cash revenue collection if at the same time spending fell well short of the target. Likewise, it is not enough to meet the spending target if spending is financed entirely through borrowing and non-cash revenue. One useful way to understand these preferences is in terms of the government's short-run and long-run concerns: the government has a short-run desire to meet its spending target, but the long-run health of the budget would require financing this spending through a greater share of cash revenue.

In this model, these preferences are characterized in terms of the government's indifference curves, shown in Figure 2 (see also Appendix Figure A2). Of course in reality the government is a composite of many competing preferences, but indifference curves are nonetheless a useful shorthand device for characterizing the trade-off faced by the government between cash revenue and total revenue. As this trade-off is modeled, the government would be indifferent to accepting lower cash revenue only if it were compensated by sufficiently higher total revenue. An important assumption is that the non-cash revenue needed to compensate the government for a lost ruble of cash revenue increases as cash revenue reaches lower levels. As a technical matter, since cash revenue is also a source of total revenue only the upper half of the figure is relevant for the model. With these preferences in mind, how tax enforcement policies of the government interact with enterprise behavior to determine cash revenue will now be described.

Cash revenue will depend on the level of cash compliance chosen by enterprises. The government can influence the level of cash compliance by threatening to seize assets of tax delinquents, although seizure can be understood to include any

Figure 2. Government Indifference Curves



action that is painful to the enterprise managers and disrupts an enterprise's operations. Strictly as a practical matter, cash revenue is defined to include any cash proceeds arising directly from the sale of seized assets. How aggressive the government is in enforcing compliance is reflected in its seizure policy, which is revealed to enterprises as a first step in the model. Enterprises then choose their optimal compliance based on this policy. The government fully anticipates enterprises' response when formulating its policy in the first step. As a final step, the government carries out its policy based on the actual compliance chosen by enterprises.

Government Pre-Commits to a Seizure Policy

The basic framework is first developed assuming that the government can pre-commit to carrying out an announced seizure policy. The main advantage to the government of pre-committing to a tougher seizure policy is that by inducing greater compliance the government reduces the number of seizures that are actually needed to meet a cash revenue target. That is to say, the policy announcement has a signal effect on taxpayer compliance. (Later I consider what happens if the government cannot pre-commit.)

In the case of pre-commitment, enterprises choose their optimal level of cash compliance in response to the announced seizure policy. On the one hand, lower cash compliance exposes enterprises to a higher risk of having their assets seized by the government, an outcome they regard as costly. On the other hand, by paying less of their tax liability in cash, enterprises that escape seizure can settle a higher share of their tax liability in kind or through mutual arrears clearing arrangements such as offsets. Consistent with the arguments made in Section IV, it is assumed that enterprises would rather settle through offsets than in cash. Based on these considerations, overall cash compliance chosen by enterprises will be greater the more aggressive the government's announced seizure policy (Appendix equation 4).

Given its preferences, the government faces a trade-off when setting its seizure policy; in response to non-payment the government can either seize assets or use unpaid tax liabilities as an imperfect substitute to cash for financing expenditure needs. One of the assumptions driving the model's results is that the amount of cash the government can get directly from seizure—for example, from the subsequent sale of assets—is less than the amount of the enterprise's unpaid tax liability. When setting its seizure policy, the government weighs the benefits of a more aggressive policy yielding greater cash proceeds—both directly from seized assets as well as from the greater cash compliance that the policy will induce—against the loss of potential non-cash revenue (Appendix equation 10).

The seizure policy the government will choose can be shown graphically by modifying the indifference curves in Figure 2. As noted above, enterprises will respond to a more aggressive seizure policy by paying more cash revenue. Since the government fully anticipates this response, its choice of seizure policy is effectively a choice of the cash revenue it wishes to target. This allows Figure 2 to be recast in terms of the trade-off between total revenue and the seizure policy, where the indifference curves of this trade-off have the same properties as the curves in Figure 2. In addition, the government faces a "budget constraint," where more

seizures decrease total revenue by reducing non-cash revenue. Since the loss of non-cash revenue is partially offset by cash proceeds arising directly from seizure, the slope of this constraint will be flatter the higher the proceeds are from seizure (Appendix Figure A3). The seizure policy chosen by the government is represented by point *a* in Figure 3.

An increase in cash proceeds arising directly from seizure would increase the efficiency of seizure. If the substitution effects of a more efficient seizure policy were greater than the income effects, the government would respond by announcing a more aggressive seizure policy. This is shown as a shift from point *a* to point *b*. Enterprises will react by increasing cash compliance.

Within this basic framework, one result of eliminating the inflation tax would be a fall in cash compliance. The logic is straightforward. Given the government's preferences, a loss of non-revenue financing will increase the government's willingness to forego cash revenue for a higher level of total revenue, effectively flattening the government's indifference curves. As a result, the government will adopt a seizure policy that targets a lower level of cash proceeds and greater total revenue, which is shown in Figure 4 as a shift from point *a* to point *b*. Enterprises, faced with a reduced risk of seizure, will lower their cash compliance, and cash revenue will decline.

Government Cannot Pre-Commit to a Seizure Policy

Given that seizing assets reduces total revenue, it is clearly in the government's interest to threaten enterprises by announcing an aggressive seizure policy and then, once enterprises have determined their cash compliance, relax the seizure policy and convert a greater share of tax arrears into non-cash revenue for financing the budget. If the government cannot pre-commit, enterprises will

Figure 3. The Government's Seizure Policy

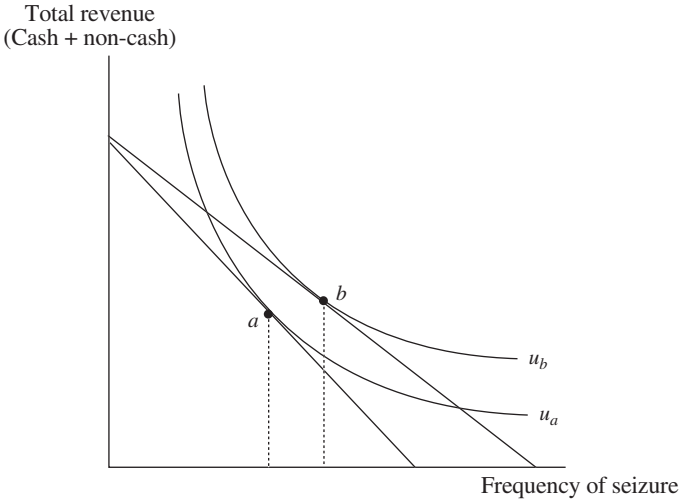
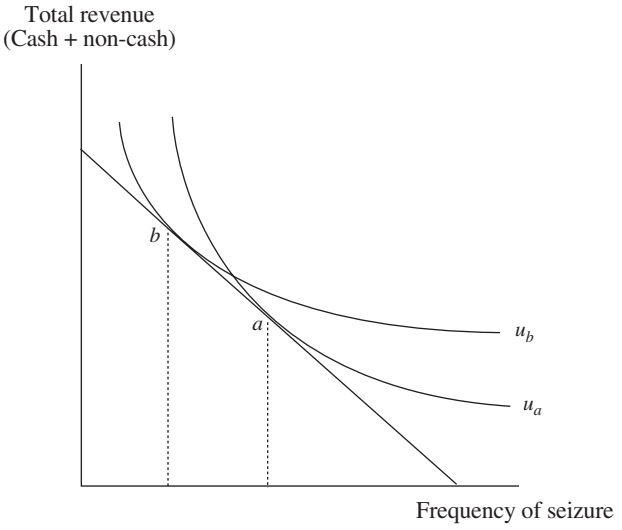


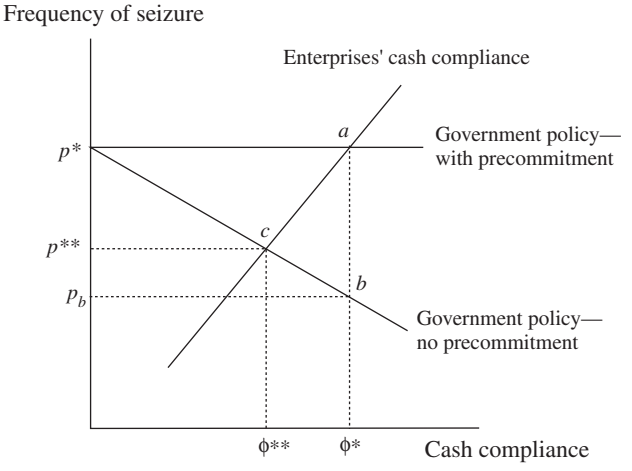
Figure 4. Decline in Non-Revenue Financing



understand that the real threat of seizure is less than the stated policy and will challenge this policy by choosing a lower initial level of compliance. I compare this result with the pre-commitment equilibrium in Figure 5.

A detailed description of the comparison is included in the Appendix, but the intuition is straightforward. Whether or not the government pre-commits, enterprises will choose compliance based on the seizure policy they expect to be carried out, with compliance increasing with a higher perceived probability of seizure. This is represented by the upward-sloping curve. In the pre-commitment case, the

Figure 5. Pre-Commitment Versus No Pre-Commitment



government chooses a seizure policy corresponding to point *a*. If the pre-commitment is credible, enterprises will respond to this policy by choosing cash compliance of ϕ^*. Ex post, however, the government would prefer to deviate from its announced seizure policy by choosing a less aggressive seizure policy corresponding to point *b.* If the pre-commitment is not credible, enterprises will anticipate the government's ex post policy preference, and will select a lower compliance corresponding to point *c*. Based on this compliance, the government will carry out seizures until the marginal benefit of the cash proceeds arising directly from sales of seized assets equals the marginal cost of lost financing (p^{**}). This equilibrium is characterized by both lower compliance and fewer seizures than in the pre-commitment case.

Tax Enforcement in a Dynamic Setting and the Low Revenue Trap

The above comparison is useful in illustrating how, in a static framework, a low return to enforcing tax compliance in the short run might cause enterprises to question the credibility of the government's enforcement policy. But is this true in a dynamic framework? That is to say, what if the government regards this as a repeated game? In that case, the government might want to carry out its pre-announced seizure policy even if this period's cash revenue benefits are very low. In a repeated game the government would take into account the consequences of its actions this period on tax compliance next period. In other words, the government might view the costs of enforcing tax compliance as a worthwhile investment in establishing its reputation, given the future benefits of a credible tax enforcement policy.

Although I have chosen for the sake of simplicity to model a one-time game, the framework can nonetheless be used to illustrate how short-run and long-run considerations might compete. If the government takes a long view, it will suffer the short-run costs of establishing a reputation. If the reverse is true, the government could find itself caught in a low cash revenue trap. As shown in the Appendix, a fall in non-revenue financing, such as through the elimination of the inflation tax, increases the premium the government places on its short-run goal of financing its target expenditures by non-cash means (Appendix, Section III). The temptation to deviate from an announced seizure policy, therefore, could increase the point that the pre-commitment itself loses credibility. At this point, enterprises would begin challenging the stated seizure policy, causing compliance to fall by much more than would be the case under pre-commitment. Given the government's need to finance short-run expenditures, the fall in compliance would further increase the short-run costs of reestablishing credibility.

This equilibrium can be characterized as a low cash revenue trap. Enterprises pay little in cash because they know the government's short-run desire to finance its expenditure needs through offsets will outweigh its long-run desire to improve cash compliance. The government, responding to low cash revenue and mounting expenditure arrears, will confirm enterprises' expectations. Attempts to break out of this trap by establishing a reputation for tax enforcement will not be credible as long as the government's expenditure needs are unchanged.

VI. Implications

Perhaps the key implication of the model is that the government can only break out of the revenue trap by cutting expenditures. In this way, the government would eliminate its reliance on non-cash revenue as a source of financing and would be free to enforce cash compliance regardless of the lost financing that enforcement would imply. Enterprises would understand this and would respond by increasing cash compliance. As the government establishes a reputation for enforcing cash compliance, cash revenue will increase and expenditures can be raised. Ironically, expenditures must first be cut to a lower level than would be the case if the government could pre-commit to a seizure policy from the beginning.

This result is closely analogous to the reduction in real balances required to reduce inflation during a stabilization. If wages and prices are set in a forward-looking manner, the government could reduce inflation without a fall in real balances if it were able to pre-commit to a reduction in the rate of money growth. The pre-commitment would translate into lower inflationary expectations, and slower wage and price growth, with real balances actually rising as the lower inflation makes money more attractive. If the policy of lower money growth is not credible, wages and prices will continue to be set based on higher expected inflation. In this case, a period of lower real balances and higher real interest rates would be required before inflation falls to the new rate of money growth. As in the case of expenditures and cash compliance, real balances must first be brought down to a level lower than is consistent with the lower inflation in the long run.

A second implication of the model is that decreasing the payoff to the government of seizure makes seizure less credible and increases the amount of seizure required to enforce a given level of compliance. If the substitution effects exceed the income effects (as shown in Figure 3), less effective seizure would compel the government to opt for a weaker policy of tax enforcement and a reduction in the frequency of seizure.

This second result is important for understanding the rise in barter and other non-cash transactions between enterprises. As discussed earlier, barter reduces the reliance on cash, making seizure a less attractive option for the government in response to non-compliance. In the context of the model, the rise in barter has the dual effect of emboldening enterprises to reduce compliance and weakening the government's determination to fight non-compliance. It is no surprise then that the sharp rise in non-cash transactions has coincided with the fall in cash compliance.[12]

As discussed, a possible alternative explanation for the revenue decline, based loosely on the Gaddy and Ickes (1998) model of the virtual economy, is that enterprises don't choose cash compliance, but pay in cash as much as they can out of their available resources. In this case, cash compliance can be increased only at the

[12]Table 5 in Hendley, Ickes, and Ryterman (1998), for example, shows a positive correlation between the severity of an enterprise's tax arrears and its reliance on barter and promissory notes in its transactions.

expense of some other cash requirement, such as wages. Supposing this to be true, one must still account for why cash revenue fell suddenly beginning in 1995. If cash revenue is limited only by the true resources available to enterprises, what would have caused these resources to have declined so suddenly?

While any number of stylized models can be contrived to link the virtual economy to the revenue decline, one plausible explanation relates to the reduction in cross-subsidization accompanying the fall in the inflation tax. Since households hold the bulk of base money, they also paid most of the inflation tax during the period of high inflation. Part of the seigniorage is likely to have been used in one form or another to subsidize loss-making enterprises, either in the form of government expenditures or through central bank directed credits to industry. In this case, the elimination of the inflation tax would represent an effective tax reduction for households, and a corresponding reduction in subsidies to loss-making enterprises. Despite the loss of subsidies, these enterprises might continue to operate as value-adding on paper through a tangle of barter arrangements and mutual non-payments. As such, their measured tax liability would remain constant, even while the real resources available to pay cash revenue has fallen.

Such an explanation certainly cannot be rejected outright and might possibly account for part of the decline in cash revenue. The key question is, which is quantitatively more important for explaining the revenue decline: (1) a reduction in cross-subsidization to loss-making enterprises; or (2) a reduction in the share of true value added paid to the government? The answer is critical for determining how to reverse the revenue decline. In the second case, as I have suggested, expenditure would need to be cut to the point that the government is not expected to rely on non-cash forms of financing, and can pursue an enforcement policy aimed at the long-run goal of raising compliance regardless of the short-run consequences. In the first case, cutting expenditure, seizing assets, and establishing a reputation would do nothing in the short run to raise cash revenue. In fact, expenditure cuts—to the extent that expenditures are a form of subsidy—might actually reduce compliance in the short run, until the loss-making enterprises are shut down and the resources they use channeled into value-adding activities.

In this regard, the case of the automobile producer AvtoVAZ is revealing. Although AvtoVAZ produces a good that enjoys a large cash market, it has for some time retained the status of the single largest tax debtor in the Russian economy. And, while the enterprise is no doubt inefficient, it would be difficult to describe it as value-subtracting. AvtoVAZ pays a small fraction of its tax liability not because there are no true profits to be made in production and sales of automobiles, but because the enterprise has not been left with any cash flow. Enterprise managers are widely believed to have used barter and money surrogates to channel all the cash flow from the manufacture and sales of automobiles to the legally separate retail arm LogoVAZ, while at the same time using transfer pricing to shift the bulk of the tax liability to AvtoVAZ. Since LogoVAZ no doubt pays close to its full share of federal tax liability, enforcing cash compliance in the current legal environment would require shutting down AvtoVAZ. The extent to which the economy is represented by enterprises such as AvtoVAZ or by, say, value-subtracting defense manufacturers is an open question.

Appendix I. A Model of Cash Compliance

This Appendix describes a simple game between the government and enterprises in which cash compliance is induced by the threat that the government will seize assets. In response to non-payment, the government chooses its optimal seizure policy balancing the revenue gained from seizing assets against the alternative of using unpaid tax liabilities as an imperfect substitute to cash for financing a chosen expenditure target. In this game, a decrease in non-revenue financing paradoxically increases the government's willingness to tolerate lower cash compliance when formulating its seizure policy. It increases the likelihood, moreover, that enterprises will challenge government's commitment to its policy of seizure, resulting in a further worsening of cash compliance.

First I consider the outcome assuming the government can pre-commit to a seizure policy. I then compare this to the case where the government cannot pre-commit, but rather adjusts its policy in response to enterprise behavior.

I. The Government Can Pre-Commit to a Seizure Policy

The game can be thought to take place sequentially. In the first step, the government determines and then pre-commits to a seizure policy. As an alternative to seizing assets, the government considers allowing enterprises to use unpaid tax liabilities as an imperfect substitute for cash to finance its expenditure target, in effect allowing enterprises to pay these taxes in kind. In the second step, each enterprise determines its optimal level of compliance so as to minimize its costs, balancing the cost of paying taxes against the likelihood that the government will seize assets if it does not pay. When formulating its seizure policy in the first step, the government anticipates the optimal response of enterprises to its policy in the second step. A key assumption of the model is that the cash value of the seized assets to the government is less in ruble terms than the value of the unpaid tax liability, if the government instead uses it to finance expenditures.

Enterprises' optimal tax compliance

To solve the model, it is useful to first consider enterprises' optimal tax compliance for a given seizure policy and then to solve for the optimal seizure policy taking into account enterprises' optimal response.

Each enterprise i chooses a level of cash compliance to minimize its costs:

$$\min_{\phi i} \phi_i t_i + v(1 - \phi_i)t_i + p_i C_i. \quad (1)$$

An enterprise's total costs equal its cash compliance (ϕ_i) times the enterprise's total tax liability (t_i), plus the cost of paying taxes in kind ($v < 1$) times the unpaid tax liability, plus the probability of seizure (p_i) multiplied by the cost to the enterprise of having its assets seized (C_i).

A crucial element of the game is how enterprises perceive the probability of seizure. One can think of the government as announcing its intention to seize

assets of a specified portion of the worst tax delinquents. After enterprises have chosen their tax compliance, the government will rank enterprises according to tax compliance plus some random adjustment factor, unknown to the enterprises, which might reflect factors such as the enterprise's lobby power, social importance, etc. The government then intends to seize assets from lowest to highest until it reaches its target number of enterprises.

Although I don't explicitly model the mechanics of such a seizure policy, I capture its essence in the following formulation. Each enterprise perceives the probability of seizure to be

$$p_i = f(p, \phi_i), \tag{2}$$

where p is the share of enterprises to be targeted by the government. An enterprise understands that the higher this share, the greater the chances that, for a given level of compliance, its assets will be seized. The function $f(p, \phi_i)$ is assumed to take a form such that the probability of seizure can be charted as Figure A1. This functional form implies that, as an enterprise's compliance falls, the chances of seizure become increasingly larger.

The enterprise solves its cost minimizing equation (1) taking into account the probability of seizure for a given government policy p. From this equation, the cost minimizing φ_i (denoted as φ_i^*) will be that which solves:

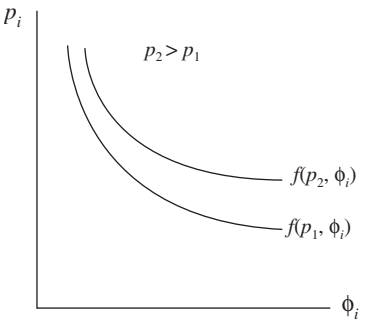
$$t_i(1 - v) = -f_2(p, \phi_i^*)\, C_i \tag{3}$$

Net marginal cost of tax payments = Marginal value of reducing the probability of seizure.

Enterprises will increase compliance until the net marginal cost of paying taxes in cash is equal to the marginal value of reducing the probability of seizure. The optimal compliance will be a function of the cost of paying taxes in kind, the government's seizure policy, and the cost of seizure.

As a simplifying convention, enterprises are assumed to be identical. This allows the industry average cash compliance to be derived directly from the enterprise's optimization in equation (3), which is expressed as follows:

Figure A1



$$\phi^* = \phi(p,\ C,\ v)$$
$$\phi_1 > 0, \phi_2 > 0,\ \phi_3 > 0, \tag{4}$$

where average compliance is an increasing function of the government's seizure activity, the costs of seizure, and the cost to the enterprise of tax payments in kind.

One noteworthy, although not surprising, implication of the enterprise's optimization is that for a given seizure policy, overall compliance will be lower to the extent that random factors, such as lobby power and social importance, factor into the government's decision to seize assets. In terms of the model, a greater randomness to the seizure process would be reflected in the limited ability of the enterprise to reduce its seizure risk by raising compliance, or in a lower value for ϕ_2 in equation (4). All else equal, this would have the effect of reducing the marginal benefit of raising compliance, and, to restore equilibrium, enterprises would reduce compliance until the marginal benefits equaled the marginal costs.

The government's optimal seizure policy

In deriving the government's optimal seizure policy, start by assuming that the government (taken here to mean the fiscal authorities) wishes to finance a given expenditure target, net of borrowing and seigniorage. This net target is assumed to be exogenous in the model. It can finance this net target either by collecting cash from enterprises or by collecting taxes in non-cash form. Non-cash tax payments can be understood to take place either directly in kind or through mutual arrears clearing arrangements such as offsets. All else equal, the government is assumed to prefer cash both because it is more fungible and because some expenditures, such as wage payments and debt service, simply cannot be paid in kind or offset.

The government's preferences, then, can be characterized by the following utility function:

$$U(E - A,\ R) \tag{5}$$

$$u_1 < 0 u_{11} < 0$$
$$u_2 > 0 u_{22} < 0, \tag{6}$$

where E is the expenditure target net of borrowing and seigniorage, which is assumed to be exogenous, R is cash revenue, and A is total cash and non-cash revenue. Based on this utility function, the government will be indifferent between cash revenue and total revenue as charted in Figure A2.

For a given expenditure target, government would be indifferent to having lower cash revenue if it had sufficiently higher total revenue. At lower levels of cash revenue, however, the amount of increased total revenue needed to leave the government indifferent is greater.

The government determines its seizure policy to maximize utility:

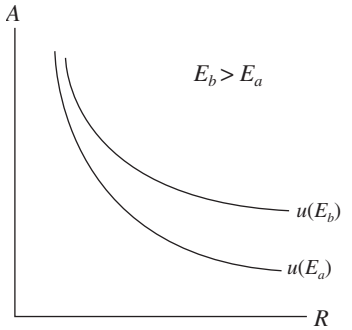
$$\max_p U(E - A,\ R) \tag{7}$$

subject to the constraints:

$$A = (1 - p)T + pQ$$
$$R = \phi T + pQ, \tag{8}$$

where p is, as before, the probability of seizure, T is the total tax liability, and Q is the cash proceeds the government recovers from seizing assets, which is assumed to be exogenous for the purposes of this model. Total revenue A is comprised of cash revenue paid voluntarily by enterprises and non-cash revenue for those enterprises from which the government chooses to accept payment in kind (the sum of which is $(1 - p)\ T$), plus cash proceeds from seizure. Cash revenue is determined by cash compliance plus cash proceeds from seizure.

Figure A2



The government sets its seizure policy anticipating the cost-minimizing reaction of enterprises in equation (4). Combining this reaction with equations (7) and (8) and, to simplify, setting the total tax liability T equal to one, the government's optimization becomes

$$\max_p U(E - (1 - p) - pQ, \phi(p,\ C,\ v) + p\ Q). \tag{9}$$

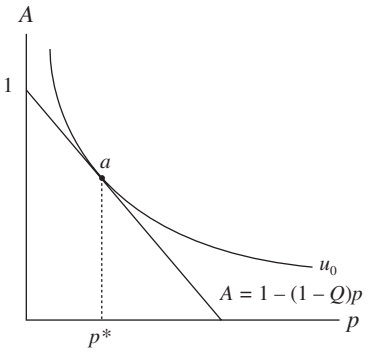
The utility maximizing p (denoted p^*) is that which solves the first-order condition:

$$u_2(p^*)(\phi_1(p^*,\ C,\ v) + Q) = -u_1(p^*)(1 - Q) \tag{10}$$

Marginal utility of cash from seizure policy	=	Marginal disutility of net loss of total revenue from seizure.

The government will set a seizure policy at that point where the marginal utility of cash, both arising directly from seized assets as well as from greater compliance, equals the marginal disutility that arises from a net loss of total revenue. The component ϕ_1—the change in compliance with respect to a change in frequency of seizure—can be understood to represent the signal effect of a stated government seizure policy on cash compliance.

Figure A3



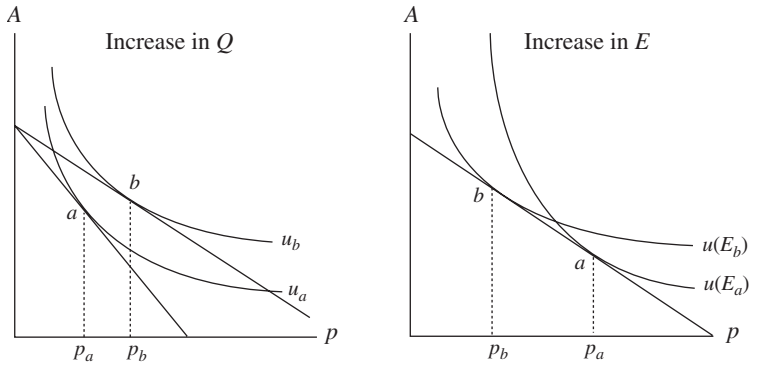
To understand the effects of the exogenous variables on the government's optimal seizure policy, it is useful to show the optimization problem graphically in terms of the tradeoff between total revenue and seizure policy p, where the policy is in effect set to target a level of cash revenue. Cash revenue R is substituted from the constraints in equation (8), combined with the enterprise reaction function in equation (4) directly into the utility function in equation (7). Utility can now be expressed as a function of total revenue A and seizure policy p and is maximized subject to the constraint on total revenue in equation (8). The optimal seizure policy p^* that solves equation (10) is shown as point a in Figure A3.

An increase in cash receipts from seizure (Q) will increase the efficiency of seizure. If the substitution effects of this efficiency gain exceed the income effects of higher cash revenue for a given seizure policy, an increase in Q will result in an increased effort by the government to seize, shown in the first panel of Figure A4 as a shift from point a to point b. In terms of equation (10), an increase in Q increases the marginal benefit of seizure while at the same time reducing the marginal cost in terms of lost total revenue. Increasing cash revenue by increasing the probability of seizure p will decrease the government's marginal utility of cash (u_2), restoring equality in equation (10). By equation (4), cash compliance will increase.

An increase in the government's expenditure target or a decrease in non-revenue financing (an increase in E) will reduce the government's willingness to seize assets. This is shown in the second panel of Figure A4. As E increases, so does the marginal disutility of the financing shortfall (u_1). To restore balance to equation (10), the government will forgo cash revenue by reducing p to the point where the higher marginal utility of cash revenue offsets the higher disutility of the financing shortfall. With a less aggressive seizure policy, cash compliance of enterprises will be correspondingly lower.

Finally, when either the cost to enterprises of seizure (C) or the cost of paying taxes in kind (ν) increases, enterprises react to a given seizure policy by increasing compliance (equation (4)). Although cash revenue unambiguously increases, the effect on seizure policy is ambiguous, depending on the impact of these changes on the signal value of seizure. On the one hand, the increase in cash revenue will reduce the marginal utility from seizure, and, all else equal, the government will

Figure A4



want to reduce seizures and increase total revenue. On the other hand, if increases in C or ν increase the *effectiveness* of seizure policy on compliance, then this will lead the government to want to increase its seizure activity. The net effect will depend on the relative size of these two effects.

II. The Government Cannot Pre-Commit to a Seizure Policy

The case where the government cannot pre-commit to a seizure policy will now be considered. In the pre-commitment case, the government first determined a seizure policy as in equation (10), and then, in the second step, enterprises determined their optimal level of compliance and paid taxes according to this level. Once enterprises paid, the government then carried out its pre-announced policy. In this sequence, however, the government's behavior is not time consistent; once enterprises chose compliance and then pay, the government no longer has an incentive to carry out its seizure policy. In terms of the model, the government's optimal p, which it determines ex ante, is greater than optimal p, which it would prefer to carry out ex post.

To see this, the government's optimal seizure policy is solved by taking the level of cash compliance as given. In this case, as before, the government solves

$$\max_p U(E - A, \phi + pQ) \tag{11}$$

subject to the constraint

$$A = 1 - (1 - Q)p \tag{12}$$

but now taking ϕ as exogenous. The optimal ex post seizure policy p (denoted p^{**}) solves the following first-order condition:

$$u_2(p^{**})Q = -(1 - Q)u_1(p^{**}). \tag{13}$$

Comparing this to the optimal seizure policy p^* given in equation (10), it is clear that since $\phi_1 > 0$ it must be the case that $p^{**} < p^*$ for any value of ϕ. Once enterprises choose their tax compliance and taxes are paid, the government can no longer influence compliance through its seizure policy, and the marginal gain to seizure is only the additional cash proceeds from seizing assets. At this point, the government will prefer to deviate from its stated seizure policy by increasing the use of tax liabilities to reduce its financing shortfall.

As can be seen from equation (14), moreover, the incentive to deviate from the stated seizure policy increases with compliance. As compliance increases, the marginal utility from cash revenue declines, inducing the government to seize less and finance more expenditures through non-cash revenue. The ex post optimal seizure policy, therefore, will be a declining function of compliance:

$$p^{**} = p(\phi) \qquad (14)$$
$$p' < 0.$$

Enterprises, of course, can be expected to take into account the government's ex post behavior when deciding their optimal cash compliance, in effect challenging the government's stated intention to seize assets. In this case, the inability of the government to pre-commit leads enterprises to lower cash compliance further.

To determine optimal compliance, enterprises now minimize costs taking into account the government's optimal ex post seizure policy determined in equation (16). As before, each enterprise minimizes

$$\min_{\phi i} \phi_i t_i + v(1 - \phi_i)t_i + p_i C_i, \qquad (15)$$

where its own probability of seizure is now expressed as

$$p_i = f(p^{**}(\phi), \phi_i). \qquad (16)$$

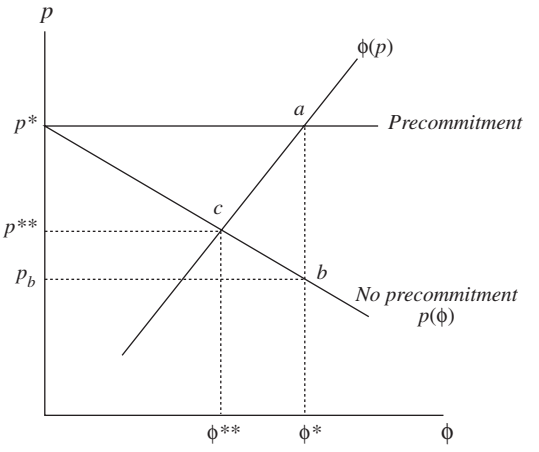
The value of ϕ_i (denoted ϕ_i^{**}) which maximizes each enterprise's utility is that which solves

$$t_i(1 - v) = -f_2(p(\phi), \phi_i^{**})C_i. \qquad (17)$$

In contrast to the pre-commitment case, enterprises now determine their own compliance based on their expectation of overall compliance. While each enterprise is assumed to perceive itself as too small for its own behavior to affect overall compliance, it nonetheless knows that all other enterprises are determining their own compliance in a similar manner. Since all enterprises are identical, the equilibrium overall compliance will be the point in equation (17) where $\phi^{**} = \phi_i^{**} = \phi$.

The resulting equilibrium can be compared to the pre-commitment equilibrium graphically. As in equation (4), enterprises will respond to a higher probability of seizure by increasing compliance. This is represented by the upward-sloping curve $\phi(p)$ in Figure A5. In the pre-commitment case, the govern-

Figure A5



ment can choose a seizure policy $p = p^*$, and the equilibrium compliance ϕ^* will occur at point *a*. Ex post, however, the government would prefer to deviate from its announced seizure policy by choosing a lower p corresponding to the downward-sloping curve $p(\phi)$ in equation (16). At the level of compliance ϕ^* the government would choose to a p corresponding to point *b*. Enterprises, of course, anticipating the government's ex post response, will select a lower compliance corresponding to point *c*, the equilibrium described by equation (17). This equilibrium is characterized by both lower compliance and fewer seizures than in the case where the government can pre-commit.

III. The Cost of Pre-Commitment

If the government views the process of seizure as a repeated game, it might be willing to endure some short-run costs of pre-commitment because it is worried about moral hazard and is seeking to establish a reputation. However, the government's cost of adhering to an announced seizure policy is itself a function of the expenditure target net of non-revenue financing. In this case, an increase in the expenditure target net of non-revenue financing might increase costs to the point that the pre-commitment itself loses credibility. As a result, compliance would fall by much more than would be the case under pre-commitment as enterprises began challenging the stated seizure policy.

From equations (10) and (15), the net marginal utility lost from adhering to a pre-committed seizure policy will be

$$-u_1(p^*)(1 - Q) - u_2(p^*)Q > 0. \tag{18}$$

As discussed above, there is an ex post incentive for the government to forgo seizure and divert the unpaid tax liability toward financing its expenditure target. If the expenditure target (E) is small, then the pre-commitment equilibrium will be

characterized by a large p^* and low marginal utility levels $-u_1(p^*)$ and $u_2(p^*)$. In this case, the marginal utility gained by deviating ex post from a seizure policy is accordingly small. If, on the other hand, E is large, the equilibrium will be characterized by a small p^* and high marginal utility levels $-u_1(p^*)$ and $u_2(p^*)$. Therefore, equation (18) shows that an increase in the expenditure target or a decrease in non-revenue financing would increase the marginal utility gained by deviating ex post from a seizure policy. This has the important implication that an increase in the expenditure target or a decrease in non-revenue financing can lead to a collapse in compliance and, at the same time, a reduction in the number of seizures, both by reducing the government's willingness to engage in a policy of seizure and by reducing the government's ex post commitment to such a policy.

REFERENCES

Citrin, Daniel A., and Ashok K. Lahiri, 1995, *Policy Experiences and Issues in the Baltics, Russia, and Other Countries of the Former Soviet Union*, IMF Occasional Paper No. 133 (Washington: International Monetary Fund).

Crotty, John, Katherine Baer, Richardson, Thomas and Richard Highfield, 1997, "Russian Federation: Addressing the Decline in Federal Cash Revenue Collections" (unpublished; Washington: Fiscal Affairs Department, International Monetary Fund).

Gaddy, Clifford G., and Barry W. Ickes, 1998, "Russia's Virtual Economy," *Foreign Affairs*, Vol. 77 (September–October), pp. 53–67.

Hendley, Kathrynn, Barry Ickes, and Randy Ryterman, 1998, "Remonetizing the Russian Economy," in *Russian Enterprise Reform: Policies to Further the Transition*, ed. by Harry Broadman, Discussion Paper No. 400 (Washington: World Bank).

Ickes, Barry W., Peter Murrell, and Randy Ryterman, 1997, "End of the Tunnel? The Effects of Financial Stabilization in Russia," *Post-Soviet Affairs*, Vol. 13 (April–June), pp. 105–33.

Interdepartmental Balancing Commission, 1997, "On the Causes of Low Tax Collection, General Causes of the Arrears Crisis, and Opportunities of Restoration of Solvency of the Russian Enterprise," Report of the Interdepartmental Balancing Commission (Moscow, Russia: Interdepartmental Balancing Commission).

———, 1998, "On the Work of the Interdepartmental Balancing Commission in 1998" (Moscow, Russia: Interdepartmental Balancing Commission).

OECD, 1997, *OECD Economic Surveys*, Russia.