CORRUPTION, BLACK MARKETS, AND THE FISCAL PROBLEM IN LDCS: SOME RECENT FINDINGS

John McLaren
Columbia University

The fiscal difficulties of LDCs, and the important role for tax evasion, are well known. Drawing on the rich literature on tax evasion in the public finance literature, much recent research has shed light on two phenomena contributing to this problem: Fiscal corruption and tax evasion through illegal transactions. This note surveys some of this recent research.

THE MEASUREMENT OF TAX EVASION

First, a note on measurement may be helpful. The extent of tax evasion in general in many LDCs is staggering, and attempts to measure it have naturally been extremely unsatisfying. The major techniques for estimation have not changed a great deal since the summary by Cowell (1990, Ch.2), and include, aside from survey methods, 'gap' approaches and monetary approaches. In the gap approach (I take the term from Alm et al., (1991)), an aggregate measure of tax obligation from the tax code and an estimate of the tax base is formed; this is compared with the actual tax receipts. Das-gupta, Lahiri and Mackherjee (1990) use this method on a time series of Indian national accounts data, and find that improving compliance measures would increase income tax receipts by at least 75 percent (ibid. Table 3). Mehta (1990) uses similar techniques to study state tax evasion in India. The monetary method begins from the presumption that illegal transactions are conducted in cash only, so a measure of cash transactions might provide an estimate of undeclared income. Making a series of assumptions about the constancy of money velocity across sectors and the initial level of undeclared income, such studies use data on aggregate cash stocks to estimate the volume of illegal transactions. An example is Polymoniouzas, Inc. (1990), which, using this technique, estimates that 25 to 60 percent of the Mexican economy is off the books. Similarly, Boguchwa and Neho (1995) estimate the fraction for Tanzania as having grown from below 10 percent to above 30 percent over the independence period.

Alm, Bahl and Murray (1991) give us what are probably the best measurements we have of tax evasion (and also legal tax avoidance) in an LDC. This study uses detailed data on tax returns and combines it with survey data in Jamaica. The authors draw two random samples of tax returns, one from self-employed taxpayers and one from pay-as-you-earn (PAYE) taxpayers, or employees subject to withholding.
The former sample is combined with survey data on the self-employed to identify the degree of non-filing among that group, and with audit data to identify the determinants and extent of underreporting in the group. The latter sample is combined with survey data to estimate the determinants of allowances as a share of total income for individual employees; these are cash supplements for specific purposes that are de jure taxable but de facto untaxed, hence a form of legal avoidance. The authors find that total revenue losses from these sources amount to 84 percent of the revenue collected, and that the majority of this revenue is lost through the very widespread non-filing of the self-employed, only a small minority of whom file any tax return at all. Further, understatement of income and the use of allowances both respond positively to the marginal tax rate, and the overall pattern of tax evasion substantially reduces the progressivity of the tax system. A study of this level of detail will not be feasible very often (partly because of the possessiveness of tax offices regarding their records), but the richness of the information revealed shows that when it is, the payoff can be enormous.

**FISCAL CORRUPTION**

There has been considerable interest recently in corruption as a problem in the provision of public goods and as a possible constraint on growth. Here we will review some recent work in the sub-branch of this area related to the tax problem.

The theoretical approaches to corruption in tax collections can be placed roughly in two categories. The more familiar one (at least in the economics literature) portrays fiscal corruption as a constraint on the government, requiring the government to consider the principal-agent problem of enforcing tax obligations through a staff of potentially corruptible inspectors. The pioneering analysis in this line by Virmani (1987) analyzed the incentive effects of various policy controls in the presence of possible corruption, and argued that the analysis of compliance incentives is greatly changed by the possibility of bribery, for example, a rise in penalties for tax evasion may easily result in a drastic drop in tax revenues, since the rise in penalty raises the benefits for the tax inspector and tax payer to incur the risks of striking a bargain. An analysis in a similar spirit arises with a simpler model in Goywami, Gang and Sanyal (1991); for example, if the number of tax inspectors willing to take bribes is increased by an increase in the tax rate, then corruption can cause a qualitative change in Laffer curve type effects, in which revenue drop when the tax rate rises. In both cases the model is set up so that these effects could not arise with uncorruptible inspectors.

Some more recent papers have emphasized the question of remuneration of tax inspectors in the presence of corruption; the discussion of corruption in practice often focuses on proposals for curing it with salary increases to the civil service. Mookherjee and Png (1995) provide insight into this problem. This paper studies the optimal incentive arrangement for a bureaucracy in which a pollution inspector must monitor a firm for compliance with pollution regulations, but the insights can be applied directly to the LDC fiscal problem. The argument that follows is valid if the word compliance is taken to mean compliance with the tax code instead of with pollution regulations. The goal of the policy in the model is to get the firm to undertake costly actions to reduce its pollution output. If the firm does not and is caught by the inspector, it is subject to a fine per unit of pollution generated. The problem for the government is that the inspector is difficult to control, along two dimensions. First, he may, on finding pollution, fail to report and offer to take a bribe instead. Second, he may simply not work very hard to find violations. The inspector can increase the probability of catching a violating firm by working harder at inspections, but this is costly to him. The government may, with some probability, catch the inspector in the act of taking a bribe, and can fire him in that case; but the government cannot observe how hard he is working, and so must structure his incentives to elicit the right amount of effort.

In this setting, the firm chooses its level of compliance with the regulation; simultaneously, the inspector chooses his level of effort; if the latter catches the former's violation, the inspector can report it, causing the firm to be fined, and receives a fraction of the fine as a commission; if it is more profitable to collude with the firm, the inspector reports no violation, and splits the savings with the firm. If collusion occurs, there is some probability that it will be discovered, an extra fine charged to the firm, and the inspector fired. It is clear in this setting that a rise in civil service compensation can have a perverse effect on civil service performance, if corruption is occurring in equilibrium. A rise in the inspector's salary is like a rise in his fine for bribery, since that is what he will lose if he is caught and fired. Thus, a rise in his salary makes it less profitable, in expectation, to take a bribe; but in an equilibrium with corruption, unless the salary goes up far enough to make the inspector give up on bribery altogether, the bribe is his return to effort; thus, his inspection effort radically drops. Paying a corrupt civil servant high wages actually induces a lower act. It may either raise or lower firm compliance; if the inspector does catch a violation, he will require a higher bribe, to compensate him for the higher risk; but because he is not trying very hard, it is less likely that he will catch a violation. The net effect on the firm's incentives are thus ambiguous.

Using similar reasoning, Mookherjee and Png show that a rise in the commission rate unambiguously increases firm compliance. Pushing this reasoning farther, the authors show that in their model the optimal policy is either a commission rate of 100 percent (or, privatization of the inspection service); or abandoning enforcement. In the case of taxation, the former option would amount to tax farming, which has its historical precedents but is not widely thought to be a realistic option because of legal problems; and the second is equivalent to shutting down the government. Neither extreme is very appealing in practice, but the point that has been made about the limits of pure wage solutions is made clearly and is an important one.

Another approach to these questions is found in Besley and McLaren (1993). That paper focuses on the optimal inspector wage problem in the presence of corruption. The problem of inspector effort is assumed away, but the government faces an additional complication: Although tax inspectors may have an incentive to collude with taxpayers to conceal taxable income, not all potential tax inspectors have the inclination. Some may well be predisposed to honesty, and unwilling to accept a bribe at any
price; ideally, the government would like to hire only these, but ex ante it can not distinguish between them and the corruptible ones. Thus, the government faces not only the moral hazard problem that is the focus of most papers in the area, but also an adverse selection problem, which can be parameterized by the fraction of potential inspectors who are corruptible in the pool from which the government hires. Call this fraction $\gamma$; it could be thought of as positively correlated with the degree of cynicism with government and negatively with a well-entrenched ethic of loyalty and honesty within the civil service.

Assume that an act of bribery will go undetected with probability $q$, and that when bribery is detected it results in firing. The parameter $q$ can be thought of as negatively correlated with the sophistication of accounting procedures and information management in the civil service. Then the optimal wage has three possible values. First, there is a minimum value for the inspector's wage, called the efficiency wage, at which it will not be worthwhile for a corruptible inspector to accept a bribe. This must be strictly above the wage the inspector could receive in his next-best alternative occupation, which we can call the reservation wage; further, the efficiency wage premium must be an increasing function of $q$. If tax inspectors are paid the efficiency wage, they all refuse bribes, and all taxes owed are collected, if they are paid the reservation wage, the corruptible ones always take bribes, but a fraction $(1 - q)$ of the time they are caught and replaced, with the evaded taxes recovered by the government. Clearly, the optimal wage would never be above the efficiency wage, or between that and the reservation wage. Finally, it is conceivable that the optimal wage could be below the reservation wage, since at least the corruptible potential inspectors are willing to work for less than their opportunity wage, knowing that they will be able to make additional income from bribery. Call the wage at which corruptible inspectors will just make their opportunity income when bribery is taken into account, the capitulation wage; if the government pays this it is giving in to the problem of bribery altogether, since it will be accepting an entirely dishonest workforce that will be accepting bribes all of the time. In this case it will collect revenues only when it catches bribery in action, or $(1 - q)$ of the time.

It turns out that each of these three wage levels can be optimal under different circumstances. If $\gamma$ is small, the reservation wage is optimal, since it would not be worth paying a premium to all inspectors just to motivate honest behaviour in a tiny minority of them, and it is not worth economizing on the wage by paying less than the reservation wage if that will convert an almost entirely honest inspector force into an entirely dishonest one. If $\gamma$ is large and $q$ is small, it is worth paying the largely corruptible workforce the small premium required. Thus, efficiency wages are optimal. However, if $\gamma$ and $q$ are both large, efficiency wages are too expensive to be attractive, and paying reservation wages would simply allow most of the inspector force to enjoy bribery rents most of the time at the government's expense. The government might as well lower their wages to reclaim those rents; thus, capitulation wages become optimal.

This model may help to explain some observations about tax administration. Wages paid to civil servants in general and to tax officers in particular well below opport-

There are a few key points about tax administration that stand out:

1. There was a very large reallocation of tax officers away from lower-yield types of audit, toward higher-yield types;
2. There was a small overall decline in yield per audit hour; and
3. This drop was the largest in those types of audits and in those regions in which the yield was the highest initially.

Although it may be possible to tell other stories as well, (i) is consistent with improved district-level managerial incentives resulting from the group portion of the RAV; (ii) is consistent with an overall rise in taxpayer compliance, resulting from a perception that tax officers now have an enhanced incentive to find (and report) violations; and (iii) is consistent with the RAV having had a stronger incentive effect in the lower yield areas due to the ceiling. Thus, the study offers intriguing (if not conclusive) evidence of incentive effects in the Brazilian case, of the sort one might hope for.

Alongside the above papers, which as noted above treat fiscal corruption as a constraint on government, a very interestingitarian strain regards such graft as a tool of government, or even, in some cases, as the point of government. An example
from the political literature is Boese (1994), who points out that in some cases a tax system with widespread evasion through graft can be useful in maintaining political control. For example, particularly in the late 1980s, graft in the Ivory Coast helped to transfer a substantial amount of income to functionaries of the ruling party through control of the bureaucracy, and in addition its very illegality was useful because it provided a threat to expose the corruption of any participant who subsequently turned out to be disloyal to the president. It thus provided, at the discretion of the ruling party, both a carrot and a stick to keep political clients in line.

In the economics literature, Marceullier and Young (1996) provide an elegant formalization of the 'predatory state', whose goal—far from maximizing the utility of a representative taxpayer—is to maximize graft. Clearly, they have in mind a kleptocracy quite protected from public pressure, and unfortunately it is not difficult to think of examples rather like that. The paradigm they use is of a government that provides order, or public services such as the enforcement of contracts and private property, which permit the formal sector of the economy to operate. This sector is taxed, which is the government's purpose in providing these services; the tax then can be appropriated by government employees through graft. The total tax is assumed to be given by tax revenues minus the cost of providing order (or, the profit of the government). Any citizen who does not wish to participate can produce output without benefit of the government's services, or in other words, in the informal sector. Such production is inherently less productive, since additional resources must be used to provide for private protection of property and similar needs; the advantage to the informal sector is that it is not taxed. Thus, there is a sense in which the government must compete with the informal sector for workers; it must wonder whether or not by charging a price too high or by providing public services insufficiently it may drive all the workers to the informal sector and have no one left to tax and extort. Thus, one may hope that the predatory inclinations of the predatory state may be substantially restrained by the presence of the informal sector.

Surprisingly, it turns out that even in competition with the informal sector, for a wide range of parameter values, the predatory government still will squeeze the formal sector mercilessly, in the sense of raising the tax rate without limit or decreasing the level of public services almost to zero. Roughly, consider two cases. If the output of the informal sector is a poor substitute for the output of the formal sector, in general equilibrium a given rise in the tax rate on the formal sector will always increase graft, because, perversely, it will increase the tax base by drawing workers in from the informal sector. This occurs because the income effect of the rise in taxes is stronger than the substitution effect; thus, the demand for both types of output falls, so the net output of both sectors must fall in equilibrium. The net output of the formal sector falls simply because of the rise in the price removed by taxes; the net output of the informal sector must fall by a movement of workers out of that sector back to the formal one. Thus, there is no Laffer curve, and the government will always wish to raise taxes a little bit more.

On the other hand, if the informal sector output is a very good substitute for the formal, unless the level of public service is quite high, workers will exit the formal sector en masse to avoid taxes, in which case the marginal benefit of order (which is proportional to the size of the formal sector) is below its marginal cost. Thus, it is not optimal to provide a moderate level of public services. The graft-maximizing decision is either enough order to create a large formal sector, or a vanishingly small level of services. In a sufficiently small economy, the former is not feasible.

Thus, in either case, if the economy is small, it falls into the 'black hole of graft', despite the presence of the informal sector. This line of argument may indeed help to explain cases like the Zairian debacle, with very low levels of public service and high levels of graft, despite the presence of the large informal sector into which any worker can exit at any time.

**SMUGGLING AND BLACK MARKETS**

In LDCs, indirect taxes are still disproportionately important in total revenue. Thus, an important form of tax evasion remains the undeclared transaction, which can take the form of smuggling to evade a trade tax, street vending or off-the-book sales to evade sales taxes, or in general operation in the 'informal', 'black', or 'second' economy.

Once again, we can think of two broad categories in which the relevant theory has been developed. In the mainstream economic approach, informal transactions have been regarded as a constraint on the government. The theory was first explored for the case of smuggling, beginning with Bhagwati and Hanssens (1978), who showed that smuggling could easily be incorporated into standard trade theory models, and if it consumes real resources, its welfare effects are ambiguous (since it can undo some of the inefficiency from a distortionary tariff, but introduces a new inefficiency by using up additional resources). This led to a rich body of literature on the positive and normative analysis of smuggling, including the widely cited Piss (1981) model of legal trade as a camouflage for illegal trade, and many others.

The other forms of indirect tax evasion have been more recently incorporated into theory. Virmani (1998), for example, incorporates an evadable output tax into a model of firm behavior. If larger firms are more likely to be audited and caught in evasion, firms will operate below efficient scale, which would not happen if the tax were not evadable. Gordon (1980) shows that in the case of a retailer with market power, formal and informal sales can coexist, because a vendor can use the option to evade the sales tax as a screening device. The point is that consumers differ in their reservation prices and also in their aversion to the embarrassment of being caught in an illegal transaction; if these two are positively correlated, the vendor can profitably offer to sell an item either for cash, thus free of tax, or formally at a higher price. The price differential exceeds the tax, so the mechanism allows the vendor to appropriate more of the consumer surplus from the high reservation price type than would be possible otherwise. Dasgupta and Gang (1996) analyze the enforcement advantages of a value added tax (VAT); in their model a VAT structure allows for complete tax enforcement with an imperfect audit system, through cross-checking of the reports of buyers and sellers of intermediate inputs. However, since this requires a sufficient
frequency of audits, which may be costly, it may be optimal to allow some evasion nonetheless, in which case the VAT will tend to distort input prices. They thus simultaneously formalize and qualify the two most common arguments for VATs in practice, their enforcement advantages and their non-distortionary treatment of intermediate inputs.

In addition, a literature has sprung up on the implications of informal transactions for optimal taxation. Bhagwati and Srinivasan (1974) studied the optimal tariff in the presence of smuggling, finding that with constant elasticity offer curves the optimal tariff is smaller with smuggling than without, but in the more general case the effect is ambiguous. More recently, attempts have been made to extend optimal indirect tax theory to settings in which informal transactions are possible.

In this vein, Kaplow (1990) studies optimal indirect taxation and enforcement in a one-good model of tax administration with and without costly evasion. He finds that the possibility of evasion implies that less public good provision, but a higher tax rate, is optimal. Lovely (1995) studies the first-order conditions for the optimal commodity tax vector in a model with costly evasion, finding that the usual Ramsey-rule logic applies without change, except that the relevant substitution elasticities must be defined in terms of legal purchases only. Lovely (1995) finds that adding Bhagwati-Hansen type smuggling to an economy in which the government chooses the optimal indirect tax vector before and after smuggling is made possible must lower welfare. Lovely (ibid.) studies the welfare effects of tightening enforcement.

McLaren (1996) also analyzes the optimal structure of indirect taxes in the presence of the possibility of black markets. In that model, it is possible for any vendor to evade sales tax, but there is always a possibility that doing so will lead to detection and a fine. The probability of detection in a given sector is determined by the number of inspectors the government has hired for that sector, times the number of transactions each inspector can audit per period, divided by the volume of transactions in that sector. Thus, the less of the good is sold, the more dangerous it is to evade the tax on it: There is a 'market thinning' effect, which can lead to multiple equilibria if the tax is high (if everyone evades, the price will be low, thus volume will be high, thus evasion will be relatively safe, and vice versa). If the government chooses its vector of tax rates and inspector forces for the various goods optimally, it will hire just enough inspectors to nullify the benefits to evaders. This allows us to calculate the required number of inspectors in each market as a function of the tax vector, and hence the enforcement cost associated with a given tax vector, allowing us to reduce the problem to the selection of a tax vector that raises the right amount of revenue while minimizing the usual deadweight loss costs plus these additional enforcement costs. If the monitoring technology is relatively effective, the latter costs will be unimportant, but with a poorly developed information infrastructure, they could be much more important than the deadweight loss.

It turns out, somewhat surprisingly, that these enforcement costs have a tendency to be concave in the tax vector. Intuitively, at the optimum, since evasion is just deterred, a high-tax sector must also be a high-risk sector for evasion. Suppose that the demand for good i is perfectly inelastic and cross-price elasticities are zero among goods with a positive tax. If evasion is just deterred, and we increase the tax on good i from $1 to $1.01, raising the number of inspectors for good i just enough so that evasion is still just deterred, then the probability of being caught at evasion must be higher after the tax increase than it was before. But then if we raise the tax once more, to $1.02, it will take a smaller increase in the good i inspection force to deter evasion: The second 1p increase in the tax is less tempting to evaders than the first, because it is associated with a higher risk to evaders than the first. Thus, because of the constraint that evasion must be just deterred at the optimum, enforcement costs have a tendency to be concave in the vector of tax rates. This implies that where these enforcement costs are most important— in an economy with poorly developed administration — the second-order condition may fail, and the optimum may take the form of maximal differentiation in tax rates across sectors. This leads to a theory of 'cash cow' taxation for poor economies, in which one sector is picked (at random, in a symmetric model) to carry the burden of all of the taxation. Only when the economy crosses a certain threshold value in its monitoring technology does it become optimal to tax all sectors. This can thus be taken to be a kind of theory of abrupt tax reforms over the course of development.

All of the above work treats informal markets as a constraint on the government. As with the case of corruption, there is also a contrarian view that treats them as a tool of government. Thus, Boone (1994) shows that official tolerance of smuggling in the 1980s by the Moomids and their allies was a way of assisting that powerful group in return for continued political support. Cross (1995, 1996) shows that the government officially tolerated the street vendors in Mexico City, who flagrantly evaded sales taxes, as part of an implicit political bargain that helped the ruling party, the PRI, keep its electoral grip on the city until the bargain came apart during the 1990s. Marcouiller and Young (1998) show that if informal goods and formal goods are complements, government tolerance of the informal sector may be useful in bolstering the official sector and thus maximizing graft.

A clever innovation in the empirics of smuggling has been presented by Decon and Aynelw (1990). They use the fact that even when data on smuggling quantities are unavailable, one can often find a reasonable estimate of the prices smugglers face. Using this price, perhaps one can estimate the implicit response of smuggling to price incentives by estimating the responses of officially recorded export quantities to the smuggling price premium. The authors do this for coffee in the case of Ethiopia, estimating smuggling prices based on the world price, the black market exchange rate, and an estimated smuggling marketing margin. They find a fairly low elasticity, 11 percent, of officially marketed quantities to the smuggling price differential, and estimate on this basis that between 1976 and 1991, 10-15 percent of total output was smuggled. Although this approach suffers from all of the difficulties faced in estimating agricultural supply functions, it may be quite useful.

Finally, a very promising approach to the estimation of welfare effects of smuggling is pioneered in Connolly, Devereux and Cortes (1998). They show that it is quite easy to use duality theory to provide second-order approximations to the effect of smuggling on national income (in the Bhagwati and Hansen (1973) sense: this is
the general equilibrium effect of allowing smuggling, holding trade policy constant. This can be calculated using data on tariffs, prices, and legal trade volumes along with estimates of the volume of smuggling. They apply the technique to Paraguay, which in the data year of 1980 had considerable import and export smuggling, and also transshipment smuggling to Brazil and Argentina in evasion of those country’s tariffs. Essentially the estimate of the welfare effect is half of the profits from transshipment (gross of smuggling costs, which are unobtainable), plus the tariff revenues on that portion of illegal transshipments imported legally, minus half of the tariff revenue lost from import and export smuggling. The authors find that national income gained more from the transshipment smuggling than it lost from the other forms, and that the net effect of smuggling was equivalent to about 2.1 percent of GDP.

As a word of conclusion, perhaps the main message of recent work is the value of a case-by-case approach in policy analysis in these areas. If smuggling, for example, could be seen as having been at some point an economic asset in Paraguay, a political tool in the Cote d’Ivoire, and a nuisance in many other places, it is unlikely that the same advice will be useful across the board. In the same spirit, the value of additional econometric case studies could be quite high.

NOTES
1. Beginning with Brunnemann (1978) and Allingham and Sandmo (1972), Cowell (1980) provides a highly readable survey. This is not meant to be an exhaustive survey. It is meant to draw attention very succinctly to some pieces of research that the last five or so years or so that I have found to be especially instructive, and especially relevant to LDCs; for the inescapably good papers that I will not have overlooked, I offer my apologies in advance.
2. The ‘reconciliation’ technique for estimating smuggling, pioneered by Bhagwati and embodied in case studies in Bhagwati (1974), might be thought of as a special case in this broad class. See Connolly, Devereux, and Curtis (1985) for a recent application.
3. Among the recent works in this broad area are Banerjee (1993), Mauro (1995), Shleifer and Vishny (1993), and Tadé (1995).
4. It increases the inspector’s incentive to monitor for a given expected firm compliance level; and for a given inspector monitoring effort and firm compliance level, it raises the expected benefit the firm would have to pay. Both effects make compliance more attractive.
5. The theory of optimal indirect taxes, as expounded by Diamond and Mirrlees (1971), and the related theory of indirect tax reform, have been broadly applied to LDC problems, most notably in Newbery and Stern (1987), and Ahmad and Stern (1991). However, results in this tradition has for the most part not paid explicit attention to sellability.
6. For example, the official prices faced by the farmers ideally would be thought of as jointly determined with the quantities; the finance of the government are surely affected by the fluctuations in the crop, and thus the implicit has rede the government monopoly on sales as a whole. However, in the absence of a model of this process, and useful instruments, the official prices are generally treated as exogenous. This possibility of source vacancy null results, such as the finding that the elasticity of coffee supply (with respect to the relative price of coffee to alternative crops) is negative.

REFERENCES
Cross, J. C. Formalizing the Informal Economy: The Case of Street Vendors in Mexico City. Working Paper, Department of Sociology, American University, Cairo, 1995.
Fighting Street Vendors Off the Street: Historical Parallel in Mexico City. Working Paper.
Department of Sociology, American University, Cairo, 1996.
Mohita, S. Evasion of State Taxes in India. New Delhi: mimeo, 1996.