PERGAMON

Social Science & Medicine 58 (2004) 237–246

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Informal payments for health care in transition economies

Tim Ensor*

International Programme Centre for Health Economics, University of York, Heslington, York YO10 5DD, UK

Abstract

There is considerable evidence that unofficial payments are deeply embedded in the markets for health care in transition countries. Numerous surveys indicate that these payments provide a significant but possibly distorting contribution to health care financing.

Unofficial payments can be characterised into three groups: cost contributions, including supplies and salaries, misuse of market position and payments for additional services. There is evidence from across the region on the presence of payment in each category although it is often difficult to distinguish between payment types.

Regulatory policy must address a number of issues. Imposing penalties may help to reduce some payments but if the system is simply unable to provide services, such sanctions will drive workers into the private sector. There appears to be some support for formalising payments in order to reduce unofficial charges although the impact must be monitored and the danger is that formal fees add to the burden of payment. Regulation might also attempt to increase the amount of competition, provide information on good performing facilities and develop the legal basis of patient rights. Ultimately, unless governments address the endemic nature of payments across all sectors, policy interventions are unlikely to be fully effective.

Keywords: Informal payments; Corruption; Unofficial Health Care Markets; Rent Seeking

Introduction

In the early stages of transition, the experience of many experts visiting former communist economies was that unofficial payments were seen by governments as a side-problem to the real issue of health sector reform. This perhaps gave the impression of an interesting but relatively rare phenomenon. Ten years on there is a growing amount of evidence that unofficial payments contribute a significant amount of revenue to the health sector. The impact of these payments is recognised to the extent that governments are now beginning to address the problem through policy. Often this policy relies on legitimising payments so that the revenue can be utilised by the public system

This paper considers the prevalence of unofficial payments in transitional economies, the impact they have on the health sector and individual access to health and the possible policy strategies that could be adopted to address them. The main focus of the paper is on the health sector. It is important to realise, however, that in most countries unofficial payments are endemic to all areas of society. As a result, the most successful strategy is likely to be one that addresses the phenomena in an intersector way.

The paper is divided into three main sections. The first examines current evidence on the prevalence of unofficial payments within a broad analytical typology based on the theory of a monopoly firm. The second section examines the policy impact of payments while the final section looks at appropriate policy responses.

Evidence on unofficial payments

The roots of unofficial activity are deep. There are at least three contributing factors. The first is that in the

*Tel.: +44-1904-433-239; fax: +44-1904-432-701.
E-mail address: te1@york.ac.uk (T. Ensor).

0277-9536/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0277-9536(03)00007-8

Soviet Union prices and outputs were strictly regulated and queuing rather than price was the normal way of rationing. As a consequence, individual connections and barter deals were of much more importance than the actual cash paid for the product. The system known as 'blat' led to the assumption that most important trades had to be supported not only by roubles but also by the promise of in-kind payment or service rendered. A teacher, for example, might tutor the children of a doctor in return for medical service. A second factor is a culture of gifts endemic at least across the Central Asian and caucus regions of the Former Soviet Union (FSU). Gifts are an expected part of professional dealings. The border between gifts and payment is a hazy one and it could even be claimed, as some philosophers have done, that identifying the person that gives the gift necessarily ensures that the relationship is contaminated and turned into a two-way trade[1]. To these historic factors must be added the general reduction in public spending on services and real-terms reduction in public sector salaries, and the increasing opportunities to earn incomes outside the public sector resulting from the break-up of the Soviet Union.

The illegal or quasi-legal status of unofficial health care payments means that establishing a strong evidence base on their prevalence is difficult. Nevertheless the last few years have seen a growing body of research literature from a diverse set of countries supporting anecdotal accounts (Thompson & Witter, 2000). The practice is seen to be so prevalent in former Eastern Europe that a summit in Luxembourg suggested that this may hold up negotiations on EU membership of Poland and Hungary (Rogers, 1999).

Rich evidence is also offered by a series of country Living Standards Measurement Surveys supported by the World Bank (Lewis, 2000). Data suggest that unofficial payments for health care are paid by a substantial proportion of patients in all FSU and Eastern European countries although between countries there is considerable variation. While 91 per cent report paying for public care in Armenia, the proportion is only 60 per cent in the Slovak Republic and 22 per cent in Albania. The variation cannot entirely be explained by the level of average income alone. Poland is one of the wealthier countries in the group yet more than 78 per cent report paying for health care.

Although there is variation, therefore, unofficial health care activity does appear to be widespread. If payments are ignored, they could have a substantial but perhaps unpredicted impact on attempts at health sector reform.

[1] For a deconstructionist view, see Keaney (1999) On the gift: a discussion between Jacques Derrida and Jean-Luc Marrion, in J. D. Caputo and M. J. Scanlon, God, the Gift and Postmodernism, Indianapolis, Indiana.

Typology of unofficial health payments

In order to examine appropriate and inappropriate policy responses to unofficial activity, it is important to be aware of the reason for, and nature of, these charges. In this section the main types of unofficial payments are examined. The analysis makes use of a conventional model of monopoly production.

Although monopoly does not characterise all health care markets, monopoly, oligopoly or monopolistic competition is important in many particularly where substantial investments as in hospital and diagnostic centre markets. Where competition is more common, the nature and level of unofficial payments are likely to change. Even within monopolistic facilities, it could be argued that since staffs often compete amongst themselves for unofficial payments, then competitive forces would have a temporising impact on the size of payments.

Contributing towards the cost of care

Most Countries undergoing economic transition have experienced a real-terms decline in government funding. The decline has been most profound in the countries of the FSU with countries experiencing declines in revenue of up to 80 per cent. Less severe reductions have also been experienced in Eastern Europe and Asian transitional countries such as Viet Nam (Witter, 1996). High economic growth in China meant that the country managed to avoid a real-terms reduction in revenue and expenditure. Spending did fall relative to GDP and there is much evidence that there has been an absolute reduction in funding for health facilities in rural areas (Ensor, 1997).

These fiscal trends were reflected in lower allocations for health and other social sectors in most countries. In general, allocations have reflected overall government revenue although the development of health insurance based on payroll contributions has reduced the impact in some countries (Ensor & Thompson, 1998). In Russia, for example, while tax revenues fell by more than 50 per cent between 1991 and 1995 collection of insurance contributions softened the overall impact on the health sector to around 28 per cent (Shiskin, 1995).

The main types of unofficial payments are conceptualised in Fig. 1. A typical inpatient institution operates in near monopoly conditions with demand for services D_1 and costs that, at some point, give rise to conventional sloping average cost (AC) and marginal cost curves (MC) as quantity increases over a certain range. In the past central planning, based on historic norms, established targets for the number of bed-days to be provided by that institution. If bed-days are taken as the main proxy for activity or volume of care, then bed-days can be represented by the line at q_1. Further, assume that

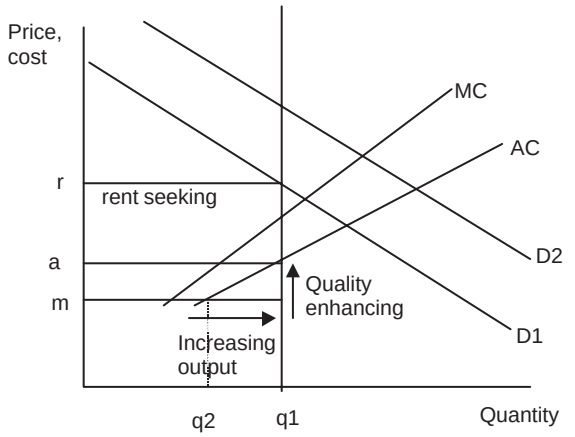

Fig. 1. Conceptualising unofficial payments.

real budget reductions mean that based on the historic number of patients treated there is only sufficient budget to part finance the real ACs. This leads to a gap between what the service actually costs (average and MC curves) and the budget per unit of output (represented by m).

The gap between resources available and those required can lead to a number of responses. One is that the hospital goes on treating the same patients but provides sub-standard care. A second response is that the hospital provides the same number of bed-days but treats fewer patients by keeping each in longer. Again quality is diminished since patients must wait longer before they receive treatment—an implicit queue. Another response is that the hospital makes an explicit reduction in the number of patients and bed-days provided (say to $q2$). In many systems, particularly post-Soviet, this has been unlikely since budget for staffing and other items is dependent on filled beds. A decrease leads to a further reduction in budget in the future.

Cost contributing unofficial payments arises when patients contribute towards the MC of their care and so close the budget-cost gap ($a - m$). These allow providers either to expand the level of activity (increasing output) or provide better services to existing patients (quality enhancing). Payments can be made to finance medical supplies through direct payments. Alternatively, patients may be required to purchase the supplies themselves from private pharmacies.

There are now numerous examples of patients that contribute towards the cost of drugs, food and other supplies either through monetary or in-kind contributions. In Kazakstan, estimates suggest that unofficial payments add at least 30–35 to public health expenditure on food and supplies alone (Ensor & Savelyeva, 1998; Sari & Langenbrunner, 2000). In neighbouring Kyrgyzstan, 90 per cent of patient reported providing food and 78 per cent providing drugs (Falkingham & Abel-Smith, 1995).

Cost contributions need not only be for material items. They could also help finance salaries. Essentially, unofficial payments might be given to ensure that staff employed in the facility reach their reservation wage—the wage which ensures retention of staff and provision of a good quality service. In theory, the issue is similar to contributions for medical supplies. Payments are made because resources are inadequate to properly finance the (staff) costs of medical treatment. Yet giving tacit acceptance to the practice of 'reasonable bribes' to medical practitioners to perform procedures that they are officially required to provide without charge is hard to accept from an ethical point of view, even if it is understandable from the point of view of personal survival.

Evidence on payments to staff is also emerging. In Albania, reports suggest that all cadres of hospital and clinic staff receive payments, the highest going to those carrying our specialist procedures such as heart surgery (Anon, 1999). In Kyrgyzstan, the same survey reported earlier also found that 25 per cent said they made a 'gift' to staff (Falkingham & Abel-Smith, 1995). In Bulgaria, Delcheva, Balabanova, and McKee (1997) suggested the unofficial cost of an operation accounted for more than 80 per cent of the average monthly wage. Further a-field, in Uganda, it is suggested that unofficial charges can double a health worker's wage (Asiime et al., 1997). A study in Poland suggested that the majority of unofficial payments are paid directly to physicians for services although it is possible that some are shared between other staff (Chawla, Berman, & Kawiorska, 1998).

The case of unofficial cost contributions is one example of a broader phenomenon of strategies employed by individuals and institutions to ensure their survival. In considering the ethics of the individual taking quite modest payments as supplement to salary, it is important to consider the alternative scenario where such payments are impossible. One possibility is that individual practitioner, to ensure survival, might leave the job for more lucrative employment. Another possibility is that hours worked are reduced to well below what is required in order that some other employment, often private medical practice, can be undertaken. Evidence on this practice is suggested in several countries. In Riga, public doctors are found to reduce their public hours in order to work in private outpatient clinics (Grieve, 1999). In the UK, hospitals have gone as far as to hire private detective to follow consultants in an effort to measure the extent of this behaviour (Rogers & Lightfoot, 1995).

Misuse of power and market position

A second type of unofficial payment is where providers exploit their power or market position in order to extract a payment for patients. In terms of

Fig. 1, payments are extracted that exceed the AC of service but are still below the individual's willingness to pay (demand curve, D_1). The maximum payment is represented by the difference between the willingness to pay and the cost contribution (r—($a-m$)). This maximum payment is likely to be reduced as the probability of being detected and punished increases.

Market power derives from a provider's monopoly or near monopoly position in the market and the principal-agent relationship that exists between practitioner and patient. There are many ways in which this could occur. One is where providers hold up treatment until payment has been made. Another is where someone with power deliberately creates a bottleneck in the provision of service, which can then be exploited by rationing that service to those that pay.

Although important for policy, differentiating bribes derived from misuse of power from those that contribute to cost is difficult. Partly it may be reflected in the magnitude of the bribe. On 11 January 1999, the Bulgarian newspaper *Standard* reported on a doctor who was arrested for taking a bribe of more than US$1100. Only a few days later, on 24 January, the same newspaper wrote that a doctor unofficially took more than US$230 (personal communication). Given that a state doctor's salary is less than US$100 per month it is hard to see such payment solely as part of a cost contribution, survival strategy.

Other examples of misuse of power are those payments taken in return for referral or for contracting. In China, for example, Bloom, Han, and Li (2000) report that providers regularly take 'kick-back' payments from pharmaceutical and other manufacturers for purchasing supplies from these companies.

Additional services

Unofficial health payments may also be made for services offered by the provider that are in addition to those promised by the state. Offering extra services generates additional demand, represented by the second demand curve $D2$, and increased opportunities for generating official or unofficial income. Patients undergoing conventional open surgery, for example, may be offered alternative minimally invasive surgery if they pay for the additional expensive supplies. Patients may also be offered better hotel services—single room, better food, air-conditioner/fan—for additional payment.

A number of studies suggest that the quality of medical and non-medical services are linked with payments. In rural Turkmenistan, it was found that in-kind payments for medical care were endemic with the quality of medical care 'tailored' to the size of payment being made (Ladbury, 1997). In Bangladesh, a complementary market for non-medical services appears to have developed among the ward boys and ayahs who make beds and clean corridors. They offer a range of additional 'hotel' services such as better food and collection of drugs from the market (Killingsworth et al., 1999). The later service is made possible by the under-funding which means that patients must obtain their own medical supplies before treatment can commence. The market is, therefore, parasitic in that it feeds off the budgetary constraints in the finance of medical services.

All unofficial payments may be made in a number of different ways. In some cases, it depends simply upon individual practitioners asking or demanding payment from a patient. In other cases, networks of staff collaborate to obtain payment. One study describes how in Kazakstan patients are referred and treated by a chain of doctors each of whom demand payment (Thompson & Rittmann, 1997). In other cases, the hospital itself may charge for services by establishing an informal tariff for services.

Policy implications

Before considering types of policy responses, it is necessary to ask whether any of this activity actually matters. The market that is created has many of the features of other private medical markets. Could such unofficial activity simply be seen as a dimension, or example, of a public–private partnership? The impact of unofficial fees can be examined from the point of view of both efficiency, equity and effective sector policy.

Impact on efficiency

The condition of many state health care organisations can be conceptualised as one of constrained optimisation. Staffing patterns based on the imposition of bed-based norms mean that organisations are not free to determine the mix of staffing and non-staffing inputs. This illustrated using a standard firm production function in Fig. 2. The budget constraint, illustrated as d–c, imposed by the state budgetary allocation to the organisation, is constrained to be spent in the way dictated by normatives. As a result the most that can be produced is X_1. If unofficial cost contributions (amounting to b–a) finance the cost of non-staff inputs then total output, as measured by the number of patients treated or treated successfully, is increased. This is illustrated as an increase in output from X_1 to X_2. From a narrow perspective of increasing output at relatively low cost, therefore, unofficial cost contributing payments might be seen to enhance technical efficiency given the constraints imposed by the budgeting process.

Yet it should also be clear that still greater efficiency would result from a more flexible approach to budget allocation. Permitting the staff and non-staff

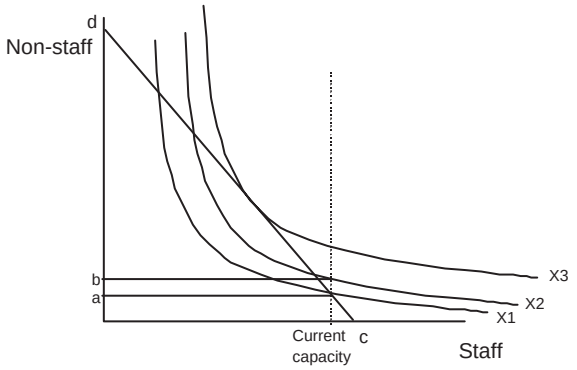

Fig. 2. Efficiency implication of unofficial payments.

budget to be used flexibly could, in this example, result in a general expansion of output to X_3 even without unofficial payments. Furthermore, it is likely that a secret system of payment is more open to abuse with, for example, payments being extracted without a guarantee that services will be improved.

Payments resulting from abuse of market power and rent extraction have no positive effect on efficiency, even the very limited kind described above. They are extracted simply because the consumer is placed in a weak market position in terms of choice of provider and immediate need of service. Rather than reducing technical bottlenecks, administrative bottlenecks are created and exploited.

Both types of unofficial payments are likely to have consequences for allocative efficiency. Cost-contributing unofficial payments imply that those that are able to provide resources for drugs and other supplies will receive service first rather than those with greatest capacity to benefit from services. Bottlenecks caused by rent-seeking opportunities restrict services to those either most able to pay or those with most influence to overcome these obstacles. Both imply that the allocation of resources is likely to be distorted away from a social optimum.

Impact on equity

Unofficial payments also have implications for equity. Imposition of unofficial payments introduces a price barrier to service in much the same way as an official user charge or privately marketed service. It might be argued that since unofficial payments are by nature flexible, this could allow providers to price discriminate and cross-subsidise so that the rich with greater ability to pay contribute more than the poor. Yet this idea is contradicted if we observe that rather than ability to pay, it is willingness to pay that determines how much is demanded. Since willingness, particularly for essential health care, is determined as much by immediate need and lack of provider choice as by income there is no guarantee that unofficial fees will extract more from the rich than the poor. Further, there is some evidence that the poor often pay more than the rich certainly as proportion of income and sometimes in absolute terms (see Ensor & San (1996) for evidence on Viet Nam and Lewis (2000) for a review of evidence across the Eastern-European/FSU region).

The issue of impact on equity, as with efficiency, is ambiguous and requires empirical verification. There may be circumstances where the social solidarity found in a small local community ensures that the rich subsidise the poor through an agreed but unofficial cross-subsidy. In other areas, where communities are less cohesive, this cross-subsidy may not exist. Intuitively, it is the elderly pensioner on a fixed income living in an urban area that may suffer more than the rural peasant living in a close knit community. This picture is not, however, supported by one recent survey that found that rural households in all income groups spend significantly more than those in urban areas (Thompson & Gorbunova, 2003).

Impact on policy

Unofficial payment may also impact on the way in which reforms of official financing incentives affect the health sector. Changes to the way in which health care providers are paid officially will, for example, have less impact than intended if unofficial payments already account for much of the income of providers.

Another side effect of unofficial payments is that the income is difficult to tax and so hinder the development of a strong fiscal base. A corollary is that high taxes are likely to make it more difficult to formalise the economy, increasing the incentive for workers to hide earnings.

Appropriate policy responses

To what extent does this unofficial activity constitute a failure of government? There are perhaps two aspects to this question. First, where government promises to provide more care than it can adequately finance. Second, through inadequate regulation to prevent corrupt practice.

Unaffordable state guarantees

One approach to reducing the funding gap is to increase the level of resources going into the health sector. Many transition countries have developed or are developing insurance schemes based on earmarked payroll taxes as a way of capturing more funding for health care (Ensor & Thompson, 1998). These have had some success in raising the level of resources, although

often not as much as expected because of the impact of recession on enterprises. A key problem is that the same macroeconomic disturbances that reduce government revenue—declining national income and break-up of large state enterprises—are likely also to impact on the ability to collect payroll taxes.

Many countries and numerous reports have discussed the objective of reducing the state guarantee. The most radical example is Georgia which suffered the largest decline in revenues and has consequently introduced sweeping reductions in state benefits. They now only guarantee to provide essential services to the most vulnerable groups. In other countries attempts to define limited 'benefit packages' have often led to little real change in what it is thought necessary to offer. Ultimately, however, reducing the size of the guarantee and also permitting providers to utilise resources more flexibly, as discussed in relation to Fig. 2, is likely to have most impact on unofficial cost contributions.

Formalising unofficial payments

If further tax or insurance revenues are not forthcoming, then a second approach is to capture out of pocket payments by formalising unofficial charges. There appears to be some support for this practice. One study in Bulgaria, for example, found that more than two-thirds were in favour of legitimising payments (Delcheva, Balabanova, & McKee, 1999). In some countries payments are already semi-formalised. China, for example, has evolved a structured system of unofficial and informal charges with white official payments—patient payments that are sanctioned by the state, grey payment—collected by the facility but not sanctioned, and black payments which are undisclosed patient bribes raised by individual practitioners (Ensor, 1997).

There is now an extensive literature on the impact of developing user charges in low- and middle-income countries (see, for example, Creese, 1991; Gilson, 1997; Wood, 1997). It is notable, however, that most articles do not examine the impact on unofficial fees. This is important since the observed impact of user charges is strongly influenced by the presence or absence of unofficial charges. If there are no unofficial payments then introducing a formal payment is likely to increase the cost of treatment—although the opportunity cost may fall if waiting times are shortened. When unofficial charges are present formal fees may substitute for the unofficial payment. The overall change in the out of pocket payment by patients may, therefore, be little affected. The net effect is an empirical question and much may depend upon the way in which charges are introduced. Charges that can be retained by health facilities to finance essential supplies and provide some incentive to health workers are much more likely to lead to a significant reduction in unofficial charges. Some studies are being conducted on the overall effect on all types of charges resulting from formalisation. More empirical work in this area is required to establish overall impact and preconditions for reducing cost-contributing unofficial charges.

It would be a mistake to think that either formalisation or restricting the scope of the state guarantee will eradicate unofficial payments. There is still the likelihood of payments associated with additional services and those arising from a corruption of power. Indeed, restricting benefits to certain groups and imposing charges on some people and services produces greater scope for deriving rents through control of quotas and access to services. There is evidence, for example, that in countries that have introduced charges, such as Viet Nam, that exemptions are often distributed to fellow workers and administrators rather than the poor (Ensor & San, 1996; Gilson, 1997). There is also evidence that exemptions used in the Soviet Union in the past for outpatient prescriptions has been similarly abused with a suggestions that pharmacists exercise some discretion over who gets exemption: those with connections benefit over those in real need (Thompson, 1996).

Regulating abuses of power

Formalisation may help to differentiate between payments that are basically survival responses from those that are based on exploitation of power by reducing the former. Eradicating or reducing the problem of unofficial cost contributions does not necessarily eliminate the rent seeking behaviour of individuals that generate unofficial income. Even if a limited package of services is provided to a restricted category of beneficiaries, practitioners can still exercise their own market position to control access in return for payment or favour.

The costs of unofficial payments are determined by the potential loss from being found out and the probability of being caught. The basis of regulation is to increase the cost of accepting unofficial payments. A penalty approach will only work if employees generally find that the penalties are severe enough to ensure that costs exceed benefits. If the wage without unofficial income is low and the main penalty is loss of job then the deterrent is unlikely to work. This strengthens the need to first address the problem of dealing with cost-contributing wage enhancements.

A number of strategies have been suggested in the regulatory literature for controlling this behaviour either by making it easier to catch offenders or increasing the penalty once caught. Baldwin and Cave (1999) summarise seven main methods (Table 1) most of which have some parallel in the health sector.

Table 1
Regulatory strategies towards unofficial payments

	Strategy	Relevance to the health sector	Comments
1	Command and control standards—backed by criminal sanctions.	Standards of behaviour/codes of conduct backed up by penalties.	Penalties only function if incentives to remain inside sector are sufficient. Regulator must be seen to be independent.
2	Self-regulation—regulation by professional association.	Medical Association enforcing standards of conduct.	Do public trust enforcers. Must be a collective incentive to enforce.
3	Incentives and market enhancing controls—general inducements for good behaviour.	Permitting private practice, internal competition.	Limited scope.
4	Disclosure—'naming and shaming' of offenders.	Public opinion surveys.	Requires viable alternative, public acceptance that payments are not beneficial.
5	Rights and liabilities laws.	Statement of patient rights that are protected under law.	Places onus of regulation on the individual, private settlement may be more attractive.

Source: Adapted from Baldwin and Cave (1999).

Command & control

One approach is to regulate through the imposition of standards backed up by criminal sanctions. For unofficial payments, this implies a code of conduct for medical employees and institutions that establish that the acceptance of such payments is not acceptable and if discovered will lead to severe penalties. The strategy is highly dependent on the enforcer of the regulations. A central pre-requisite is that the regulator should be independent from the industry it seeks to regulate. This is particularly important in countries where there is a lack of trust in public administration. In transitional countries, there have been a number of corruption scandals involving health authorities. In Russia, the head of a regional health insurance fund was dismissed and later prosecuted for fraud including mis-use of insurance revenues for the benefit of fund officials (Tchugaev, 1996). In Macedonia, the health minister was sacked over corruption claims including diversion of humanitarian aid bound for Albania (Makedonija, 1996).

Self-regulation

An alternative to command and control regulation is to devolve the task of regulation to a professional body. It is common, for example, in a number of countries for quasi-autonomous medical councils (such as the General Medical Council in the United Kingdom) composed of representatives of the medical profession and other representatives to maintain professional standards and adjudicate over matters of mis-conduct. Could such bodies be used to reduce the prevalence of unofficial payments?

As with command and control systems, it is important that professionals have an incentive to respond to possible penalties. With professional regulation, the incentive must go further so that there is a general incentive for the profession as a whole, as represented by its council, to minimise unofficial payments. Another important pre-requisite is that it is not only the professions that are represented on the council. Lay-representation can help to ensure that vested interests do not prevent regulation from being effective.

Incentives and market-enhancing controls

Much regulation is concerned with isolating individual occurrences of delinquent behaviour. Another approach is to develop competitive incentives and disincentives that are known to have an impact on aggregate behaviour. Their advantage is that they do not require policy makers to spend resources on finding those that are infringing the regulation. Some disadvantages are that they can be costly if introduced for a large sub-group and can be a blunt instrument since they penalise or reward all no matter whether behaviour changes.

Legalising private practice by doctors working within the public sector is one way of expanding competition. This can increase the choice available to patients. It may also open up a new source of income to public practitioners. If this source is seen as complementary, with public practice by building a reputation for private practice, this may itself increase the benefit of a public job and reduce the propensity to take unofficial payments. On the other hand, encouraging private practice can mean that practitioners are tempted to spend less time fulfilling their own public sector commitments. There is already evidence that public practitioners in other countries develop their private practice through referrals from their public clinics (Lewis, LaForgia, & Sulvetta, 1996; Nandakumar, Reich, Chawla, Berman, & Yip, 2000).

Competition from the private sector does not, however, address the issue of the unequal competition arising from free point-of-delivery services. Staffs still charge unofficial fees provided that prices to patients are

lower than the private sector levels. Reducing the scope for payments requires extending the principle of competition to employees within the public facility. If there are few good doctors for patients to choose from, the potential for unofficial payments is greater. Feeley et al. suggest that unofficial payments are in fact lower in Moscow despite higher levels of professional service and generally higher income levels (Feeley & Sheiman, 1999). One reason for this may be that the greater concentration of specialists lead to more competitive conditions within and between public facilities. If this is so, then regulation should be targeted at facilities that offer high value service but where both internal and external competition is weak.

Disclosure

The aim of regulation through disclosure is to put pressure on offenders by naming the worst and the best. Policy-makers might name hospitals that came out best in terms of quality standards in the hope that those lower down the list will be pressured into improving standards. Likewise it may be possible to produce a similar list of institutions taking unofficial payments. Official information on payments is, of course, scarce but it might be possible for opinion pollsters to survey patients in an effort to find out how much is paid, for what and to whom. In Kazakhstan, for example, the popular newspaper Caravan has produced surveys on unofficial payments in hospitals (Caravan, 1996). Information available may help to reduce the prevalence of payments.

There are problems with this approach. One is that, to act as a disincentive, payments must be seen by consumers to be unambiguously a bad thing. If payments are seen as a way to improve access to good doctors or improve the efficiency of production then far from being seen in a negative light, institutions that offer unofficial services may actually be preferred. The information collected will have to reflect the rental or excessive element of payments that do not add to the quality of services.

Rights and liabilities

A clear code of consumers rights protected in law may help to reduce the demand for unofficial payments. This is dependent upon a functioning and fair legal system that can be used to protect the rights of consumers through criminal prosecution or civil actions. Many transitional countries cannot depend upon such systems either because they are themselves corrupted or because they are not independent of the state. In Kazakhstan, a key informant suggested that only one person had taken a hospital court for eliciting informal payments. The provider later got off by paying a bribe to court officials. In these circumstances, it is hard to have much trust in the development of consumer rights although at least one research organisation in Moscow is developing a system of public liability insurance.

Endemic corruption and spill-over effects on society

Institutional corruption, of the type mentioned above, raises the general question of whether significant regulation of unofficial health care payments is really possible given the climate of endemic corruption prevalent in most of the transitional economies. Transparency International ranks all countries in a corruption perceptions index according to surveys of business experiences of dealings (International, 1999). The Heritage Foundation has developed a black market index again based on perception surveys (Heritage_Foundation and Wall_Street_Journal, 1999). Each of these indices suggest that transitional countries of the FSU and Eastern Europe have an extremely high level of corruption and black market behaviour. Only Estonia is seen to be comparable with Northern European countries.

In countries where corruption is endemic, an approach that concentrates on sanctions taken against individual offenders for taking unofficial payments within the health sector is unlikely to have much impact. Corruption is part of daily life to the point at which it is no longer considered illegitimate. In this case, approaches to unofficial fees for health care must be considered in the broader context of regulation and incentives that attempt to reduce the overall level of corruption in society.

Conclusion

There is growing evidence to suggest that unofficial health care fees are likely to distort health care priorities and change the impact of health system reform. It is important that ministries of health and international organisations take into account this impact in designing more effective policy.

The distinction between the issue of cost-contributing unofficial payments and those based on abuse of power, authority and market position is a key first stage in unravelling the impact of unofficial fees and beginning to design effective regulation. Formalisation of some unofficial fees, with careful monitoring of their impact, may help to reduce their prevalence. Beyond this the development of a clear system of patient rights, simple procedures for complaints, transparent contracts of employment and targeting those facilities in which collusion among professionals leads to 'networked' health care fees are all important strategies. More work is required to test out the impact of combination strategies on the prevalence of fees.

Fundamental to any policy response to unofficial charges is a greater understanding of society wide corruption. Unofficial fees is not a local problem to be addressed by the sector ministry. It is important that the problem is tackled at the national and even international level. Currently, although there is a growing literature on strategies to correct corrupt behaviour (for example, archives of the World Bank and Transparency International), little of it is health sector specific. This means that it does not address the nature of much of health care corruption such as the essential nature of health care, difference between corruption and survival strategies and impact of the principle-agent relationship.

References

Anon, 1999. *Paying for health care*. Sofia, Bulgaria: Klan.

Asiime, D., McPake, B., Mwesigye, F., Ofumbi, M., Ortenblad, L., Streefland, P., & Turinde, A. (1997). The private-sector activities of public-sector health workers in Uganda. In S. Bennett, B. McPake, & A. Mills (Eds.), *Private health providers in developing countries*. London: Zed Books.

Baldwin, R., & Cave, M. (1999). *Understanding regulationTheory, strategy and practice, Oxford: Oxford University Press.*

Bloom, G., Han, L., & Li, X. (2000). *How health workers earn a living in China*. Brighton: Institute for Development Studies, IDS Working Paper 108, p. 24.

Caravan, 1996. *If you don't grease, you won't go*. Caravan newspaper.

Chawla, M., Berman, P., & Kawiorska, D. (1998). Financing health services in Poland: new evidence on private expenditures. *Health Economics*, *7*(4), 337–346.

Creese, A. (1991). User charges for health care: A review of recent experience. *Health Policy and Planning*, *6*, 309–319.

Delcheva, E., Balabanova, D., & McKee, M. (1997). Under-the-counter payments for health care: Evidence from Bulgaria. *Health policy*, *42*, 89–100.

Delcheva, E., Balabanova, D., & McKee, M. (1999). *Unofficial fees in Bulgaria*. Unpublished.

Ensor, T. (1997). What role for state health care in Asian Transition Economies? *Health Economics*, *6*(5), 445–454.

Ensor, T., & San, P. B. (1996). Access and payment for health by the poor in Northern Vietnam. *International Journal of Health Planning and Management*, *11*(1), 69–84.

Ensor, T., & Savelyeva, L. (1998). Informal payments for health care in the Former Soviet Union: Some evidence from Kazakstan and an emerging research Agenda. *Health Policy and Planning*, *13*(1), 41–49.

Ensor, T., & Thompson, R. (1998). Health insurance as a catalyst to change in former communist countries? *Health Policy*, *43*(3), 203–218.

Falkingham, J., & Abel-Smith, B. (1995). *Financing services in the Kyrgyz Republic: The extent of private payments*. London: Report to Know How Fund.

Feeley, F.G., Sheiman, I.M., & Shiskin, S.V. (1999). *Health sector informal payments in Russia*. Unpublished paper, Moscow.

Gilson, L. (1997). The lessons of user fee experience in Africa. *Health Policy and Planning*, *12*(4), 273–285.

Grieve, R. (1999). *Unofficial payments in Riga, E-mail correspondence*. London: United Dental and Medical Schools.

Heritage_Foundation and Wall_Street_Journal, 1999. 1999 *Index of Economic Freedom*, http://www.heritage.org/index/execsum.html.

International, T., 1999. *The Corruption Perceptions Index*, http://www.transparency.de/documents/cpi/index.html.

Keaney, R. (1999). On the gift: A discussion between Jacques Derrida and Jean-Luc Marrion. In J. D. Caputo, & M. J. Scanlon (Eds.), *God, the Goft and Postmodernism*. Indiana: Indianapolis.

Killingsworth, J., Hossain, N., Hedrick-Wong, Y., Thomas, S. D., Rahman, A., & Begum, T. (1999). Unofficial fees in Bangladesh: Price, equity and institutional issues. *Health Policy and Planning*, *14*(2), 152–163.

Ladbury, S. (1997). *Turkmenistan health project social assessment study*. Report to the World Bank.

Lewis, M. (2000). *Who is paying for health care in Eastern Europe and Central Asia, Washington*. World Bank, unpublished monograph.

Lewis, M. A., LaForgia, G. M., & Sulvetta, M. B. (1996). Measuring public hospital costs: Empirical evidence from the Dominican Republic. *Social Science & Medicine*, *43*(2), 221–234.

Makedonija, N. (1996). 27 February.

Nandakumar, A. K., Reich, M. R., Chawla, M., Berman, P., & Yip, W. (2000). Health reform for children: The Egyptian experience with school health insurance. *Health Policy*, *50*(3), 155–170.

Rogers, A. (1999). Health-care bribes in states aspiring to join the EU. *The Lancet*, *350*(9093), 1758.

Rogers, L., & Lightfoot, L. (1995). *NHS hires private detectives to spy on doctors*. London: The Sunday Times.

Sari, N., & Langenbrunner, J. (2000). *Out-of-pocket payments in health care: Evidence from Kazakhstan*. London: Euro-Health.

Shiskin, S. (1995). *Metamorphoses of the Russian health care reform*. Moscow: Voprosy Economici.

Tchugaev, S. (1996). Chevrolet, for health and wealth. *Izvestia*, *16*(5), 129.

Thompson, R. (1996). Reimbursement policy for pharmaceuticals in the Russian Federation: An overview. *Maxwell Stamp report to EU TACIS Russia Pharmaceutical Sector Development Project*.

Thompson, R., & Gorbunova, L. (2003). *Health care access, utilisation and spending in Kazakstan: Findings from a household survey*, in preparation.

Thompson, R., & Rittmann, J. (1997). A review of specialty provision: Urology services. In R. Thompson, T. Ensor, & J. Rittmann (Eds.), *Health care reform in Kazakstan, compendium of papers prepared for the World Bank Health Reform Technical Assistance Project*, 1995–1996.

Thompson, R., & Witter, S. (2000). Informal payments transitional economies: Implications for health sector

reform. *International Journal of Health Planning and Management*, *15*(3), 169–187.

Witter, S. (1996). Doi moi and health: The effect of economic reforms on the health system in Vietnam. *International Journal of Health Planning and Management*, *11*, 159–172.

Wood, A. (1997). Annoted bibliography on user charges in developing countries. *ODA research programme: Improved care of diseases of childhood*, http://cich.ich.ucl.ac.uk, http://cich.ich.ucl.ac.uk.