

DRUG PRICE CONTROLS: Regulations Without A Cause

W. Duncan Reekie
School of Economic and Business Sciences
University of the Witwatersrand
Johannesburg 2050
reekiew@sebs.wits.ac.za

1. INTRODUCTION

Concerns over the prices of, and levels of expenditure on pharmaceuticals have been expressed in many countries and for several decades. By contrast patients demand new and ever more effective drugs and governments provide patent protection and hence price freedom as an incentive to and reward for innovation. This paper looks at some aspects of this apparent policy dilemma – price control and price freedom – as it affects drug markets in both developed and developing countries. In particular I shall draw on UK and SA experiences.

This paper is divided into four main sections. First, the UK Pharmaceutical Price Regulation Scheme (PPRS) is described briefly. It is remarked that the PPRS itself is only one of a network of interlocking and, on occasion, interdependent controls on expenditure. It is also noted that when individual prescribers (or small groups of prescribers) in the UK were given direct incentives to control expenditures and select less expensive products, they did so to a greater extent than when reliance was placed exclusively on the blanket controls of the PPRS.

Second, the theory of optimal pricing is outlined. The discussion suggests that neither controls nor markets are likely to arrive at optimal price levels. However it is theoretically probable that market-determined prices are spontaneously more likely to be pitched at optimal levels than are controlled prices substantiate this.

Third, price controls in general tend to encourage shortages. Empirical evidence tends to the laws of demand state that, if prices are held artificially low, more of a given product will be consumed, and, over time, this change in consumption will be even greater. There is nothing special about pharmaceuticals to suggest that these laws are in suspense. Indeed two factors suggest the reverse. When individuals have to pay (a prescription drug charge) the effect is to discourage expenditures. If they receive medicines free, however, consumption rises. In the longer run the impact of underpricing may be even more insidious. If the price of health care (including pharmaceuticals) is held artificially low then, by extension, so too is the price-or cost-of ill-health. The consumption of ill-health will therefore rise. Or, to put it differently, holding down pharmaceutical prices artificially, and providing them at low or zero cost to patients, may well encourage moral hazard in lifestyle choice. People, after all, are free to choose less healthy lifestyles, and if the cost of ill health is held down that choice is encouraged.

Fourth, we have the issue of patents. According to the *World Development Report*, 'intellectual property rights are important for encouraging innovation,

particularly in such areas as medicine and agriculture. When creators of knowledge do not retain exclusive rights of ownership for a period of time, there is far less incentive to produce new knowledge.' The problem then is that the presence of patents can hamper distribution of goods to poorer consumers who 'can seldom afford the prices charged by patent owners'. So 'developing countries have responded. . . by proposing safeguards [such as] ensuring of access to essential medicines at reasonable cost'. But if this implies lower prices for poorer markets today it can, through arbitrage, translate into lower prices elsewhere and so less innovation for all markets tomorrow. Commissioner Lamy of the European Commission stated that then 'there must be barriers to re-export' to wealthier markets.

However, there are strong arguments which suggest that holding down prices reduces incentives for investment in the production of pharmaceuticals. In the longer term investment in research and development (R&D) will also be adversely affected by distorted incentives. The paper concludes by suggesting that the administrative costs of the PPRS and associated controls be reduced or abolished, for the benefit of consumers, producers and tax-payers. Government rather should concentrate on that area of the market where the need is greatest—namely assisting the indigent or the chronically sick to pay the market price for medication. Hardship can thus be alleviated by efficiently targeting limited government resources, and the industry and patients will be encouraged to invest in and consume pharmaceuticals at levels appropriate for the community's needs.

2. HOW PHARMACEUTICAL PRICES ARE CONTROLLED

The UK Case

Why is this so? The alternative to regulating prices by diktat is to control them by competition. Competition, if deemed to be absent, may have to be supplemented or replaced by controls. This is particularly so if there are unusual market conditions in the demand and supply of a product which inhibit the full workings of the competitive process. This much seems logical. But if the controls themselves, and the administrative apparatus which accompanies them, are unnecessary (because the market is already competitive, and because the demand and supply conditions necessary are in fact present) then the outcome will not be an improvement. Rather there will be a suppression of the competitive process and a distorted result which may well have long-term harmful and unintended consequences for producers and consumers alike.

The British government does not set pharmaceutical prices as such. Rather, monopsonistic (single buyer) of drugs through the National Health Service

(NHS) it attempts to control expenditures on pharmaceuticals in a variety of ways. First, consumption by patients is affected by the prescription charge. Second, it influence prescribing patterns by doctors by laying down guidelines and also rules on what can and cannot, and on what should and should not, be prescribed. Finally, and this is the issue which we examine first, it controls profits directly at manufacturer level.

Profit Controls

Pharmaceutical manufacturers supplying the NHS are subject to the Pharmaceutical Price Regulation Scheme which is a non-statutory agreement between the industry and the Department of Health. Despite its name, the scheme aims to control profits rather than prices. A company is assigned a rate of profit which takes account of its particular circumstances, especially its economic contribution, realized and potential, in terms of capital investment, research expenditures, employment and foreign earnings. The scheme, the origins of which may be traced back to 1957, has most recently been and unbranded generic renegotiated in a revised version which came into effect in 1999. Essentially the PPRS limits the amount of profit a company can earn on its total sales to the NHS. If profits rise above a specified figure, product price reductions may be permitted. The initial price of a new product is not controlled.

The scheme began in 1957 as the Voluntary Price Regulation Scheme (VPRS). Prices to be paid by the NIS were determined by reference to export prices, or, where appropriate, the equivalent of the generic price, or according to a formula taking into account manufacturing, distribution and marketing costs. In 1969 companies were required to submit annual financial returns, to enable the government to judge whether 'excess' profits were being made.

In the face of rising inflation and falling profits the VPRS was eased in 1972, but in 1973 the Monopolies Commission investigated the prices charged by Roche for *Librium* and *Valium*, and ultimately concluded that, given their market share the company was making excess profits.

Against this background, in July 1976 the Secretary of State introduced controls to cut expenditure on promotion from 14 percent of sales to the NHS in 1975 to 10 per cent by 1979. There was no ban on expenditure above 10 per cent, but it was not counted as an allowable expense in the calculation of profit from sales to the NHS. In 1986 this figure was reduced to nine percent of total sales to the NHS and companies were 'fined' the whole of any excess in spending above 9 per cent, thus effectively trebling the cost of the marginal pound spent on promotion. In 1993 the promotional levy or fine was removed but expenditure above a

company's limit was not an allowable expense. Also, in 1986, unbranded generics were excluded from the scheme. This was because the 'amounts reimbursed to pharmacists for generics are based on the level of prices set competitively in the market'.

In 1978 the scheme was renamed more fully to reflect its nature: the VPRS became the PPRS. Further, the 1978 changes incorporated a more explicit system of rewards for companies which contributed to the economy by investment or exports.

The return on capital earned by companies on sales of NHS medicines is matter for negotiation within a published range between companies and the Department of Health. Since 1988 the industry and government have tried to remove the target rate of return from the debating arena by agreeing on a figure which will take account of 'any relevant and significant alterations since the last PPRS rate change in the underlying average return on capital of British industry as brought out for example in relevant changes in the FT 500 index or of any relevant policies that may be generally in force on the appropriate levels of profitability on public sector business'. Whether movements in some average accounting return on capital in British industry should be 'reasonably' correlated with that of the pharmaceutical industry begs the question.

In addition, companies may in certain circumstances be permitted to retain profit above their target range. This discretionary allowance (previously called the 'grey area' and in later versions the 'margin of tolerance') has varied. Under the 1986 scheme companies could retain profits if they did not exceed 50 per cent of the agreed targets. The Department could allow such retention of profits if they arose from: 'the launch of a new product, improved efficiency or other factors clearly arising from the own company's efforts'. Under the 1993 scheme, companies had an automatic right to retain profits within the margin of tolerance of 25 per cent above or below target.

As a consequence of negotiations leading up to the 1993 Scheme, an immediate general across-the-board price reduction of 2.5 per cent on all branded products was implemented, followed by three years of price restraint. In 1999 the reduction was 4.5 percent.

Other Methods of Controlling Expenditure

In addition to profit controls, the NHS uses other devices to limit expenditure on pharmaceuticals. The impact on firms, like price controls, is to restrict revenues.

The methods have included cash limits indicative prescribing, limited lists, encouragement of generic prescribing, the use of prescription charges and, as mentioned earlier, the controls of advertising volumes through the PPRS, as well as competition policy.

One successful method, but abandoned after a change of government, was cash limits. These applied to fundholding general practitioners. There were some 35,600 GPs or family doctors in the UK providing a prescribing service for their patients. After the introduction of the fundholding scheme in 1991 the number of fundholders grew as follows:

Number of Fundholders

Year	GPs	Practices
1991	1,801	308
1992	3,364	621
1993	7,197	1,355
1994	10,201	2,362
1995	15,407	4,257

Source: Compendium of Health Statistics, Office of Health Economics, 1997.

These practices covered over 50 per cent of the population and it was presumed that this percentage would have increased. Fundholding practices had somewhat different prescribing incentives from the conventional practice in the NHS. Fundholders received cash-limited budgets which could be allocated for certain defined purchases (including pharmaceuticals) on behalf of their patients. The budgets were approved and distributed directly and annually by the more than 100 integrated Health Authorities.

A fundholder who overshot his cash limits could not continue to supply unless other monetary arrangements were made with NHS approval. Alternatively at year end any cash surplus could be appropriated by the fundholder for approved practice purposes. The incentives were therefore not to overspend but to conserve resources and utilise the cash budget cost-effectively.

Pharmaceuticals represented only one item in the fundholder's portfolio of expenditure items. The verdict on how the system affected prescribing behaviour was still only partially in when it was abolished. However, the Audit Commission indicated that the first results were encouraging. During 1992/93, for example, fundholders achieved average expenditure levels on pharmaceuticals 9.4 per cent less per surveyed prescribing unit than those of other practices. Further, the growth rate of expenditures by fundholding

practices ranged from 2.0 to 4.3 per cent less per annum than the 12 per cent rate of increase of non-fundholders. In its final report on the topic to the commission provided more data confirming this evidence. In 1993/94 expenditures on prescribing were less on average among all fundholders vis-à-vis non-fundholders, but only with the earliest practices to opt for fundholding (Wave 1) was the difference statistically significant. Wave 1 fundholders would, presumably, be the most commercially aggressive groups since they were the first to self-select into the system. On the other hand diminishing returns from savings apparently set in for the Wave 1 group. Fundholders of all vintages again had a rate of increase in prescribing expenditures below the non-fundholding average, but only among the last group (Wave 3) was that slower rate of increase statistically significant. The law of diminishing returns has not been repealed. Whether its impact had been felt surprisingly quickly-or not-is another issue. The fundholder, after all, was not only under an incentive to conserve costs but, like any other pre-payment system; there was an incentive to underspend. Only when the cash nexus is direct between patient and provider is it possible in principle to avoid moral hazard of this sort altogether. The prescription charge of £5 .65 from April 1 997 did not achieve this cash nexus even for pharmaceuticals since nearly 84 per cent of scripts written were exempt from even this modest co-payment.

The Health Authorities also monitor GP prescribing and are required to initiate early and effective action regarding any 'excessive prescribing', defined as significant divergence of actual expenditures from that planned or predicted. GPs are expected to stay within a 'target budget'. The Audit Commission found that 85 per cent of practices (92 per cent of larger practices) overspent these amounts in 1991/92. The total overshoot was 7.5 per cent of the budgeted amount. In setting indicative amounts Authorities consider practices' historic spending patterns, comparable average costs for the district, the special circumstances of the practice including high-cost patients, anticipated changes in demand and an allowance for the forecast increase in the cost of drugs. (For the now defunct fundholders, the prescribing costs were actual allocations financed from the overall drug budgets allocated to the NHS regional office.)

Corrective actions, such as withholding of remuneration, may be taken against prescribers who persistently overspend. Since August 1988 the prescribing patterns and cost of every GP have been monitored by an information system known as PACT (Prescription Analysis and Cost). From 1991 each GP received a monthly budget statement of practice expenditure and indicative amounts. In addition prescribing costs of a practice compared with local and national patterns and with practice data from the previous year are provided. The availability of drugs where generic prescribing would reduce costs is also indicated.

Another route followed by government to exert downward pressure on pharmaceutical expenditures is the Limited or Selected List. First introduced in April 1985, this meant that fewer drugs were available on prescription under the NHS. Specifically seven drug categories were mentioned: minor painkillers, cough and cold remedies, laxatives, indigestion remedies, vitamins and tonics and sedatives. A selected list of around 400 products in these categories was duly excluded. The objectives were to encourage purchase from chemists over-the-counter and without prescription. The 'blacklisting' was not unsuccessful as measured by some yardsticks. For example, according to the Office of Health Economics, expectorants and cough suppressants-the second most commonly prescribed sub-grouping in 1970 had fallen to the nineteenth ranking by 1990. A further ten categories were announced in late 1992, some more products have subsequently been blacklisted and in other cases product prices have been reduced in order to remove the threat of listing. The Selected List thus not only proscribes prescription, but removes the previously available blanket pricing 'freedom' under the PPRS for the unlisted, and so tacitly approved, products in the seventeen categories concerned.

Government has also devoted substantial efforts towards promoting greater generic prescribing. As a result the proportion of prescriptions dispensed in generic forms grew appreciably to account for over 40 per cent of the total number of prescriptions dispensed by chemists, as compared with 25 per cent in 1985. The Department of Health actively encourages dispensing chemists to purchase parallel imports rather than domestically produced equivalents. From 1991 the chemist has been allowed to retain the full difference between the official list or reimbursement price and the price he paid for the parallel product. If the individual pharmacist fails to react to this incentive he will lose some of it anyway when the Department of Health 'claws it back' in subsequent remuneration awards made 12-24 months later. These are computed on the assumption that the parallel imports observed in aggregate by government through its licensing and statistical trade records are distributed evenly through the retail industry as a whole.

A rationale for any price-control scheme such as the PPRS and the associated expenditure controls is that not only may UK prices be too high but that unregulated market forces may fail to bring about price levels that would result in reasonable profits in drugs. This belief may exist because on the demand side of the market doctors and patients do not pay directly for the medicines they prescribe and consume and are therefore indifferent to the prices paid by the NHS; or it may be because on the supply side firms are believed to have the ability persistently to price their products above reasonable levels because competitive market rivalry is absent, for whatever reason.

But are competitive pressures absent on the supply side of the market? The following sections provide empirical and theoretical reasons why they may, in fact, be in place. The trends in innovative pharmaceutical price movements are what economic theory would predict. Moreover, it will be demonstrated that attempting to control these trends is likely to be counter-productive, both in the short run and, more damagingly for health, the demand side-also in the long run.

3. THE SIMPLE THEORY AND PRACTICE OF OPTIMAL PHARMACEUTICAL PRICING

Conceptually optimal price-regulation is straightforward. Prices should simply be set at levels where they equal marginal cost. In practice this is difficult since 'costs' are essentially subjective phenomena. If only the decision-taking manager knows-or believes he knows-what the alternatives are which are foregone in setting a particular price, there is no means of some outsider, such as a regulatory authority, second-guessing what is only known to the decision-taker. The problem is compounded, even conceptually, when joint costs such as unassignable overheads (or R&D) have to be included somehow in the essentially subjective marginal costs.

Even if the essential subjectivity of costs is ignored, a major remaining difficulty for policy makers is the confusion caused by the words 'perfect competition'. These suggest that the phenomenon (and its associated condition of price equal to marginal cost) is ideal. In fact, the concept is merely a predictor. And in an innovative industry, where unit variable costs are relatively small (as in drugs) but fixed costs such as R&D expenditure are relatively high, the problem of correctly defining marginal cost (even if it can be presumed to be objective) is well nigh insuperable. Unless this is understood, we are in the hypothetical elementary textbook abstraction of homogeneous products, where no improvement or innovation can bring extra profits exclusively to the supplier. If he cannot reap the rewards there is then no incentive to make improvements. Hence 'perfect' price competition (or price equal to short-run marginal cost) is a prescription for the non-existence of innovation.

The basic problem is that the price of a drug must exceed its marginal cost of production in order to contain its own assignable overhead costs and also that portion of common costs allocated to it, both apportioned over the quantity of the product sold. How great should this excess be? And how should it vary drug by drug? The theory of price discrimination helps provide an answer.

What is the ability to price-discriminate? The word is technical, and has no pejorative meaning. It simply indicates that a product is being sold at different

(non-marginal cost-related) prices. This can be justified on welfare ground in that it enables a firm to appropriate to itself 'consumers' surplus' and thus increase producer profits. This is economically desirable where marginal production costs are very low (as in pharmaceuticals) or close to zero (as in providing for the crossing of a river on an already constructed bridge).

Consider Figure 1. DE represents the demand curve for a medicine. Assume zero marginal production cost. There are three different market segments with differing levels of willingness or ability to pay. These different segments each have the same volume potential, say a million units. The profit-maximizing price and output level for the firm is to sell 2Q units at £0.5 per unit. Sales and profits (as there are no costs) would be equal to the area ABCO. Provided this area exceeded the expected innovative costs of product development, the firm would conduct R&D and produce and sell the product as described. Consumers would also receive, 'free' as it were, welfare benefits equal to the consumers' surplus (triangle ABD).

There are two welfare defects in this situation, however. First, if the innovative costs exceed ABCO (i.e. £1 million) the product would not be developed at all, despite the fact that at output level C total welfare (producer's revenue plus consumers' surplus) equals ODBC, i.e. £1.5 million. Second, total possible gross consumer welfare is not restricted to ODBC; it is equal to the whole area under the demand curve (ODE), i.e. £2 million. A single uniform price of £0.5 would result in a 'deadweight-loss' of BCE. But no firm would produce three million units to avoid this. To sell these extra units, the firm would have to reduce the price to £0.25, and it would realize an income of only £0.75 million.

The firm would invest up to £2 million in R&D, however, if it could practise price discrimination. This is because it would sell one million units to the least price-sensitive segment at £0.75, a similar number to the next segment at £0.5 and another one million units at £0.25 to the most price-conscious market. It would then earn revenues equal to the shaded area in the figure, a number approaching £2 million.

In short, price discrimination enables the firm to service people who otherwise could not afford to purchase. It enables it to expand its output beyond the physical level it would select if limited to choosing a uniform, profit-maximizing price; and as the discriminatory alternative is more profitable, it brings into the firm's realm of choice R&D projects that it would otherwise not consider.

The British Office of Fair Trading describes the process as follows:

There are many areas of business where (price discrimination) is a usual and legitimate commercial practice. For example . . . in industries where there are large fixed costs and low marginal costs (the cost of supplying each additional unit of output is very small compared to the initial investment to set up the business). In most markets undertakings are normally expected to set prices equal to their marginal cost but in industries with high fixed costs an undertaking which did so might never be able to recover its fixed costs. It may therefore be more efficient to set higher prices to customers with a higher willingness to pay. In general price discrimination will not be an abuse in such industries if it leads to high levels of output than an undertaking could achieve by charging every customer the same price.

Thus price discrimination benefits all. Poorer people less able or unable to pay the normal, uniform profit-maximizing price gain access to a product that they would not otherwise have. Today's medicines, for example, can be made available more cheaply. Producers reap greater profits, which give an incentive to engage in further research in order to develop tomorrow's medicines more quickly. And a portion of these additional profits comes from the better-off, who have the most obvious desire to purchase innovations (as indicated by their willingness to pay) and who tend (sometimes, but not always) to have altruistic feelings towards the poor and less privileged.

The ability to practise price discrimination depends, of course, on the preservation of market segments as distinct markets. This requires, in markets subject to innovation, the presence of a degree of monopoly protected by patent or other legal arrangements. In our simple example, with zero marginal costs the price would be driven to zero by competition in an unprotected environment. And it would be driven to zero in each segment. By extension, the ability to practise discriminatory pricing also depends on a lack of arbitrage or leakage between segments. A firm can charge different prices in the segments only if it is not possible for a third party to come along and buy cheap in one segment and sell dear in another (but at a lower price than the existing firm is currently charging).

The truth of the matter is that patent protection combined with price discrimination enables higher rates of economic development through the encouragement of an R&D-based industry as well as low prices for essential medicines. But the twist in the tail for developing countries attempting to achieve these dual benefits is the threat of parallel imports. Without some means of preventing parallel imports, it will not be possible to price-discriminate effectively.

The figure above can also be interpreted dynamically. The different segments can represent the total market at different points in time, with prices falling as period succeeds period. Revenues are greatest in the middle period, but higher prices can be obtained early, and lower prices charged later to maintain sales, resulting in greater total revenues over time than if a fixed price had been charged throughout.

A major advantage of the PPRS as a price-control mechanism is that it permits such spontaneous price differences over time, albeit with the constraint of a return on capital adjusted, as discussed, for a range of factors such as export potential and innovation encouragement. The relative prices will be closer to optimality than uniform mark-up pricing.

The Empirical Evidence

The large volume of available empirical evidence supports the above theory. For example, this author looked at the daily dosage pricing of all NCEs launched in each of the USA, UK and Holland relative to daily dosage costs of existing products in the various therapeutic sub-markets. Major NCEs (as rated by governmental or independent authorities) tended to be introduced at either relatively high prices (often a multiple of, existing prices) or, on occasion, at discounted prices if succeeding rivals were anticipated. The less important (incrementally) the NCEs were, the lower were their relative prices, until, at a 'me-too' level, discounts became common. Further, over time the higher priced innovations fell in price.

A more recent study I carried out repeated and reinforced these conclusions for all types of new pharmaceutical products, not just NCEs. Most new products, imitations, new presentations or combinations, and new indications tend to be introduced at substantial discounts, pulling down the average price paid in any given therapeutic market. This continues until new blockbusters pull up average prices paid once again and the cycle then repeats itself.

Thus on the supply side pharmaceutical prices appear to result from understandable and justifiable competitive pressures. But can the same be said for the demand side of the market? The next section attempts to answer this question.

3. THE PERVERSE INCENTIVES OF DEMAND SIDE PRESSURES ON PHARMACEUTICAL PRICES

On the supply side of the market pharmaceutical prices are under competitive pressures. The evidence supports this assertion, so too does the theory.

Unfortunately, the very existence of price controls (whether of the PPRS variety or reference pricing or product-by-product controls as exist in some European countries) carries with it the implication that prices are not competitively determined. Certainly the PPRS awards price freedom subject to an overall profit cap. But, as noted, the outcome of that is to distort the incentives for the supplying industry (and so overly encourage or discourage the innovative activity wanted by consumers). If the cap is too low, innovative investment may be discouraged. The cap may be too high. Recall that the profit cap was arrived at after allowing deductions of specified promotional and R&D costs. If the allowable costs are 'too high' managers and investors may be prompted to invest in pharmaceutical innovation in the UK when, in an unregulated market, alternative investment opportunities either elsewhere or in other industries could provide greater value to consumers. In the 1990s the allowable R&D charge under the PPRS around 20 per cent of sales compared to an industry average was spend world-wide of 15.3 per cent.

The problem, of course, is insoluble for policy makers. The worth of an existing product or of an innovation (not the costs of production or innovation, however defined) cannot be judged by whether it provides the firm with a return on capital equivalent to the 'FT 500' (or any index) but rather what it is worth to the patient in terms of what he or she will willingly pay to acquire its benefits. That can only be determined in a free marketplace. Indeed the generation of such information is a prime purpose of markets as institutions. It cannot be determined by a third party in Whitehall, although a proxy value can be imposed. The essence is individual choice varying with individual circumstances and knowledge. In practice some, but only some, choosers (patients) can make such a choice individually. In others, agents (prescribers and reimbursers) make the choice on behalf of patients. A plurality of agents working for different principals is more likely to arrive at a choice close to that of non-homogeneous individuals than is some monolithic government agency.

Price studies can only show that competition is working. They cannot evaluate in advance the incremental benefits which the consumer has not yet judged. This is not to argue, of course, that there should be no pressure for cost containment. The issue is how that pressure should be applied. This returns us full circle. The alternative to regulation is competition to enable consumers to assess information about product attributes and, in particular, price; and on product alternatives and, in particular, alternative prices. Prices facilitate informed choice.

'Customers' as decision-makers in the health care market include the patient, the doctor-prescriber, and often but not always, a third-party reimbursers.

How can government strengthen the demand side further to foster competition? How can it encourage an informed environment rather than a directed one? How can patients and prescribers be empowered so that their individual wishes and wants are responded to by suppliers? Cost-conscious consumers will demand cost-effective products and treatments. But what is cost-effective in their eyes, need not necessarily be so in the eyes of a reimbursers.

The tried and tested way, of course, is price, freely paid, not merely freely set (as it is, subject to profit capping, under the PPRS).

Because prices are not freely paid in the NHS at least two perverse outcomes occur. These are short-term and long-term moral hazard. Both increase total expenditures on pharmaceuticals, the very opposite of what the PPRS aims to achieve.

Short-term Moral Hazard

Moral hazard is the tendency for insurance itself to increase the likelihood of the insured event. It is a well-discussed topic. People either take fewer precautions against the insured incident or are more willing to consume the service provided by the insurer (i.e. there are both long- and short-term effects). Supply and demand in the health-care market do not meet in a cash nexus. Physicians (other than fundholders) have little incentive to conserve prescription costs. Patients receive prescriptions free or at a nominal charge and so have no direct expenses and little incentive to refuse medication (at whatever the cost to the NHS or an insurer) or to avoid initiating it in the first instance.

But does the United Kingdom health-care market have to incur the short-run moral hazard costs that result from pricing pharmaceuticals at a near-zero level? Certainly imposing a co-payment is a common way to design an insurance contract to attenuate moral hazard. But few prescriptions bear a co-payment. The fundamental question is whether or not pharmaceuticals are insurable goods at all.

The two basic principles of insurance are that occurrence of the event insured against should be rare and that the financial consequences of the event insured against should be large in relation to the individual's income or wealth. People are willing to pool their financial resources in such circumstances in order to shift

the financial risk of the event to the insurer. They are willing to pay what may be termed, from their perspective, a risk premium. That is the amount over and above the expected loss required to administer the risk pool, make payouts and assume an aggregate risk.

When the event insured against does not meet these two criteria the risk premium will be so large in relation to the expected loss that an insurance market will not come into being. Thus fire, accident and theft insurance markets exist. A car is usually and easily insured against a crash by its owner. Damage to its bodywork could cost several hundreds of pounds to repair but this, given the country's population of cars, is an event which occurs relatively rarely. Conversely insurance markets tend not to exist for punctures. The transaction costs of organising an insurance market for punctures and of making frequent and small payouts to claimants would be such that the premiums would be so high relative to the repair costs that insufficient demand would exist to make it worthwhile even for technically and administratively efficient insurers to do so. If such a service were provided at zero price by government, of course, motorists who suffered a puncture would undoubtedly claim. Indeed the moral hazard effect (namely clients fortuitously or opportunistically taking less care to avoid the event) would result in higher levels of punctures reported as, for example, drivers took less care over rough ground or those with worn tyres deliberately induced punctures to gain the insurance benefits.

Comprehensive coverage for pharmaceuticals does not meet the criteria for insurable goods. The reasons are obvious. The products are not consumed on an infrequent and unpredictable basis nor do they form a large part of the income or wealth of consumers. Only 9.7 prescriptions were written per head of population in 2000 at an average total cost (net ingredient cost plus distribution channel costs) of £10.55 per script or £113 per head per annum.

This is not to say that some pharmaceutical consumption would not be regarded as 'naturally insurable'. For example the very poor might regard an expected annual expenditure of £113 as a large part of their wealth, or certain chronic conditions requiring high absolute expenditures may occur and persist, thus raising annual or lifetime expenditures per head well above that statistical average (e.g. diabetes). At the time of writing, however, 85 per cent of prescriptions written are dispensed free of the £6 prescription charge, and this exemption was granted, according to the Office of Health Economics, to 'over half of the UK population' in 2000.

A third, and not unimportant, reason for believing that pharmaceuticals in general are not appropriate products for insurance is that good risk management attempts to minimise the administrative costs of insurance. This requires that the

insured are not burdened with absurdly high premiums (or taxes in the case of the NHS) in order to 'pay for the processing and monitoring of large volumes of low-value claims. We have already seen that pharmaceutical prescriptions are written and dispensed frequently and their cost (to the NETS and net of administrative expenses) is low. Yet the costs of administering the PPRS and its complementary regulations, together with the existing system of chemist reimbursement by the Department of Health (with its elaborate formulae, including 'claw backs' on notional parallel imports), is, at the very least, elaborate.

Long-term Moral Hazard

The two laws of demand state that, as price falls, quantity demanded will rise; secondly, in the longer run, the relationship will be even stronger as people adjust their general lifestyles further to take advantage of lower prices. (For example, if petrol prices fall, people will drive further and more frequently, so consuming more fuel. In the longer term they will also buy larger vehicles which are less efficient in their usage of petrol.) The markets in health care or in pharmaceuticals are not exempt from those laws. This is why some medical insurers reward their clients for staying healthy, by offering discounts on gym memberships and other bonuses for maintaining healthy life-styles.

Pharmaceuticals were initially regarded by the NHS as part of the insured package (or more precisely as part of the tax-funded package-since the NHS is mainly funded from central taxation not from the nominal National Insurance Contributions of employee and employer). In other words, the NHS and its pharmaceutical coverage is a type of social insurance, not casualty insurance.

Moral hazard, short-run or long-run, is an inevitable function of insurance. Epstein, however, argues that the long-run hazard costs are greater with social insurance. He contrasts casualty insurance with social insurance. Casualty insurance allows people to exchange a stream of premium payments for an increase in aggregate utility. The utility increase comes from the smoothing out of their flow of wealth, irrespective of whether their health status is good or bad. The market is stable because the insurance transaction provides anticipated gains to each insured person. No-one will exit. But the market for casualty insurance provides no redistribution in income or wealth, and coverage will not be complete. As in any competitive market, there will be exclusions of those unwilling to pay the market price. In insurance markets higher risks require greater premiums. Social insurance can, conversely, accomplish redistribution, and does so by mandating inclusion of the entire population in the risk pool.

Cross-subsidisation across groups occurs, and long-term moral hazard problems arise. Epstein states:

...some reduction in overall physical fitness will follow from . . . subsidies of health care. As the price of ill-health goes down, the willingness of individuals to take health risks will increase.

Redistribution is thus achieved at the cost of mandating inclusion of all and of increases in moral hazard.

So market-style attempts to attenuate short-term moral hazard problems such as prescription charges for all (or indeed full cost recovery) can be thwarted by regulatory intervention. The problems of long-term moral hazard in social insurance such as the Nil-IS are especially disturbing. A report in *The Economist* stated:

NHS doctors and hospitals may be under strain in some poorer areas . . . largely because the health of the local inhabitants is poorer, not because the NHS is spending less per head there.

Certainly it is cheaper for the poor and unemployed to queue than it is for high-income earners. But the report went on to indicate that it is not simply reported sickness that was higher.

So too were 'objective' indicators such as infant mortality, likelihood of death in middle age and susceptibility to diseases such as cancer and heart conditions. Such indicators have been diverging further from their respective UK means than they did in 1948, and in fact have been dropping below their absolute values of 50 years ago. They are measures of health that are not susceptible to short-run, but long-term, moral hazard. The poor eat less fresh fruit and vegetables and more junk food. They take less exercise, are less health-conscious and smoke significantly more. The UK, through the NHS, has encouraged moral hazard in lifestyle choice. And the incentives for such behavior are signaled very clearly by the free provision of most pharmaceuticals.

Suppression of innovation in Cost Containment

In few industries has the distribution process for a final product been subject to potentially more effective and innovative methods of organisation in recent years than in pharmaceuticals. The industry is in a state of flux and cost-containment measures have developed in ways which, only a decade ago, would have been unimaginable to the outsider. Pharmaceutical benefit managers (PBMs) have

emerged in several national markets which monitor prescription, consumption and usage of medicines. The monitoring may be pre- or post-consumption; it may be linked to formularies, or to total disease management programmes (TDMs). The range of innovation is enormous.

Yet, if markets are precluded from working by price controls, such innovation may be suppressed or never uncovered. The price-control effect may be implicit. The inability of a price to rise because of regulation may discourage the search for an improvement. Or the effect may be explicit. Thus, as an example of the latter, in the UK an Executive Letter (EL(94)94) was sent from the NHS Executive Headquarters on 8 December 1994 (for expiry only on 8 June 1 996) instructing all tiers of NHS management responsible for pharmaceutical purchases down to GP fundholders that they 'must not make commitments to purchase drugs' on 'preferential' terms from 'companies offering disease management packages'. Such packages are part of the PBM programmes which in recent years have resulted in downward pressure on final pharmaceutical prices elsewhere in the world. As the Department of Health's 1996 *Report to Parliament* pointed out, these activities '*reduce the effective price paid by payer*' (emphasis added).

The reasons given by the NHS for the ban on distribution channel innovations in the UK include the detrimental 'bypassing of community pharmacy' and, presumably with unconscious irony, the 'undermining' of the information required by government to implement its Pharmaceutical Price Regulation Scheme (PPRS). In mid-May 1996 EL(94)94 was superseded by a new consultation document. It allowed pilot schemes to go ahead, but it was still highly cautious in tone and did not encourage disease management.

4. PATENTS

Economic development requires the protection of private property and the application of the rule of law. As even these minimal requirements are lacking in many developing countries, perhaps it seems rather irrelevant to discuss the protection of intellectual property (IP). After all, if the average citizen is unable to protect even his or her own land or to enforce contracts in a court of law, of what use is the hypothetical ability to protect intellectual property? Actually, it turns out to be of much use. A recent survey found that many inventors from developing countries patent their inventions in the United States. This suggests that there is considerable demand for patent protection in those countries. If they had a cheap and reliable system of patent protection, it seems reasonable to suppose that there would be more invention. The benefits for the wider

community in developing countries would be twofold: a more vibrant local industry and products that are more relevant to the desires of local people.

Table 1: Per capita GNP, R&D employment, technology exports and patents filed: international comparisons

Country	Per Capita GNP (\$) (1999)	R&D Scientists and Engineers per million people	High-technology exports as percentage manufacturing exports (1998)	Patents filed (1997)	
				Residents	Non-residents
Brazil	4,420	168	9	2,655*	31,947
India	450	149	5	10,155	6,632*
South Africa	3,160	1,031	9	n.a.	n.a.
United States	30,600	3,676	33	125,808	110,884

*For 1996

Source. World Development Report, 2000-2001 (Washington, DC: World Bank, 2000).

Moreover, better patent protection would encourage innovators from developed countries to engage in joint projects. Rather than surviving on the leftovers from the research and of development (R&D) of developed countries and producing chemicals and pharmaceuticals invented by others, firms in developing countries would become not only inventors but also the legitimate partners of developed country inventors, engaging in complementary R&D.

Countries such as Brazil have in their hinterlands large numbers of as yet unscreened plants, vegetables and other sources of substances that may have therapeutic potential. There are stocks of untapped knowledge about some of these substances existing in the minds of 'traditional healers' and in oral folklore. But without incentives to capitalize on these assets, they may remain untapped or – particularly in the case of folk traditions – be lost altogether if the small, local and often tribal communities in which they are embedded become submerged in the larger, anonymous nation-state.

South Africa has a similar potential, along with large existing clinical research assets of both physical and human capital. Pharmaceutical firms have for decades carried out disproportionate amounts of clinical trial activity there. The country has a large tertiary medical sector and clinical access to both First World and Third World disease. Yet another country with complementary R&D potential is India. Sharing some of the characteristics of Brazil and South Africa, it also has a large, sophisticated and well-developed manufacturing sector that could readily exploit chemical and pharmaceutical inventions and discoveries.

In sum, patent protection in developing countries would lead to a larger R&D-based local industry, which in turn would lead to economic growth. The Table provides some basic data comparing patents and technology in Brazil, South Africa and India with similar figures for the United States.

Critics argue that, as with intellectual property in developed countries, introducing stronger IP protection in developing countries is likely to lead to higher prices for the protected products. This concern is heightened by the importance of some of those products to people in developing countries, specifically the drugs used to combat deadly diseases such as malaria and AIDS. However, the reality is that if patent-holders are able to price-discriminate, they will be willing to sell those drugs at whatever price the buyers are willing to pay – as long as their costs are covered. Moreover, as we know in general, without patent protection, potential inventors have less incentive to develop their ideas.

Preventing parallel imports

Parallel trade is the exporting or importing of a product through channels other than those authorized by the owner of a patent. It occurs if no compulsory licence is granted in the country of importation but imports continue all the same. One remedy open to the patent-owner is to ask the government in the importing country to enforce the patent in the import market. Another, less clear-cut remedy is for the patent-owner in the exporting country (often the same firm or a firm associated with the one being commercially damaged by the imports) to challenge the right of the exporter to resell to the importing country. (The exporter will presumably have obtained legitimate legal title to the goods through, for example, their purchase in the exporting country.) This involves challenging the principle that the product, if legally acquired by the exporter in the country of origin, can be legally resold in the importing country. This belief, that international resale is legitimate everywhere and always, is the doctrine of international patent exhaustion. And here the World Trade Organizations's (WTO) Agreement on the Trade-Related Aspects of Intellectual Property Rights

(TRIPS) is of no assistance. Article 6 of TRIPS states that for 'the purpose of dispute settlement . . . nothing in this Agreement shall be used to address the exhaustion of intellectual property rights'.

The issue is contentious and unresolved. Some argue that parallel importation is a normal competitive practice arbitraging away price differentials and therefore that governments in importing countries should ignore the practice. They argue that local laws, in either the importing or the exporting state, which deny the principle of exhaustion, are equivalent to protectionist trade barriers. This argument can be supported by the 'first-sale doctrine', according to which once ownership has been transferred, the patent-holder has already received full value for the patent. Resale is of no concern to it.

The patent-holder has 'alienated' or 'exhausted' its exclusive rights to control distribution of the product once it has placed it in the stream of commerce. The counter-arguments are, *inter alia*, that Article 28 of TRIPS specifically recognizes the right of patent-holders to prevent the unauthorized sale of both domestically produced and imported versions of their products. A bar on parallel imports of identical products acquired from a sister company is simply a logical application of the patent-holder's right to be the exclusive importer or domestic producer.

Economists cannot adjudicate on these legal issues. They can, however, argue about when and why conditions exist that should or should not favour price arbitrage. Conversely, they can argue about when and why immediate, as opposed to gradual, convergence of price on marginal cost is or is not optimal. For example, we know that much of the sound and fury in the debate on parallel trade of medicines within the European Union (EU) is misplaced. The arbitrage process there – trading within the EU – is due in the first place to artificially induced price differences brought about by varying national price control regimes existing side-by-side within a legally defined single international market.

Parallel trade within the EU is thus easy to explain, if not necessarily to condone. It is perverse to have different price controls within one (allegedly single European) market. It is the individual controls resulting from each member state wishing to direct its own social and health policies that are the problem, not the pricing behaviour of the manufacturers or the marketing behaviour of the wholesalers doing the international trading. An economist, then, would condemn not parallel trade within Europe but the price controls that give rise artificially to arbitrage.

Parallel importation or trade in the rest of the world is more difficult. Most countries are economically segmented from one another in a way that member

states are not in the politically determined, so-called common market of the EU. The market conditions that affect price determination do indeed vary between non-EU countries. However, and this applies to European markets too, it can be difficult initially to understand that market conditions also vary within countries.

In South Africa, for example, the government is essentially a monopsony (single purchaser for the bulk of the pharmaceuticals market (some two-thirds by volume but only one-quarter by value). Patent-holders practise price discrimination within the total national market in order to recover the proportion of development overhead attributable to that market mainly from the much wealthier private sector. Only patent protection, coupled with the identifiability of more than one market segment (government and private) with different price sensitivities, enables this to be done. In short, the absence of parallel importation, combined with patent protection, facilitates disproportionately large purchases by the government on behalf of poorer members of the population. Table 2 shows South Africa's per capita income relative to other countries. It shows how South Africa can be regarded commercially as having two distinct market segments, gauged by income and thus presumably by ability to pay.

Table 2: Per Capita GNP at purchasing power parity, 1999: an international comparison

Country	\$
South Africa	27,699*
United States	30,600
High-income countries	24,430
Middle-income countries	4,880
Low-income countries	1,790

*Top 20 per cent of income-earners.

Source. World Development Report, 2000/2001.

One example of successful price discrimination of this sort is for bronchodilators. The public sector benefits quite clearly from domestic price discrimination.

About 80 percent of South Africa's population relies on (mostly unpriced) care through the government sector, while the remaining 20 per cent relies on a private sector system much like that of the United States. The price discrimination between these sectors works to the advantage of low-income patients. Prices are higher in the low-volume private sector and lower in the high-volume public sector. Government purchases account for 66 per cent of industry volume but only one quarter of revenues, while the private sector generates volumes and values of turnover close to a reversal of these proportions. For example, in the late 1 990s the South African private sector paid R28.99 for the asthma inhalant Ventolin, while the state sector paid R5.66.

Table 3: Selected drugs prices (R) in the South African and other markets, 1997

Product	Ventolin (out of patent)	Zantac (still in patent)
SA Private sector	28.99	205
SA Government sector	5.66	28
SA weighted price (by volume)	9.05	109.97
United Kingdom Price	15.73	169.34
World Average Price	22.86	161.04
Product sourced via IDA*	8.45	On patent (n.a.)
Lowest-priced source identified by IDPIC**	8.34	On patent (n.a.)

*International Dispensary Association

**International Drug Price Indicator Guide

Source: Company sources and price lists cited in Reekie (2000)

The price discrimination system that has evolved spontaneously in South Africa (a middle-income country) illustrates the theories discussed earlier. Price discrimination depends on patent protection and on the absence of illegal imitation, of parallel importation from countries where third parties can buy and import to the higher-priced segment from abroad, and of resale from the state sector into the private market. These conditions have been met. The South African example above shows how patents, properly protected from parallel imports, can simultaneously help to alleviate poverty, and provide revenues above manufacturing costs to fund R&D.

5. Conclusions

Despite varying regulatory frameworks, competitive pressures tend to be present and working. Furthermore these market pressures are indeed competitive and not monopolistic. The results of studies outlined above indicated that pharmaceutical firms do behave in a manner which tends towards competitively determined prices. Two factors, however, patents and third party reimbursement produce countervailing tendencies. The success of the now defunct fundholding experiment in the UK and the increasing use of personal medical savings accounts elsewhere, suggests that prospective reimbursement schemes are more successful in holding down pharmaceutical expenditures than are retrospective payment systems. This in turn could impact on the direction of R&D. Retrospective payment schemes, other things equal will encourage quality increments (generally cost increasing) while retrospective systems will encourage innovative competition directed at treatment cost reductions.

Neither of these alternatives is necessarily what patients would choose in an uninsured market place where a proper cash nexus exists. There is some doubt as to whether pharmaceuticals should be regarded at all as insurable goods. But as long as private health insurance is either subsidised through the tax system (employees have their premiums paid net of personal income tax) or through some system of social insurance then the demand for insurance and hence for pharmaceutical coverage will be higher than otherwise. The impact on global R&D of the proposed extension of Medicare cover for seniors in the USA can only be guessed at, but its direction and magnitude will surely be suboptimal.

Rather, along with reducing controls on the supply side, government should be encouraging competitive forces by fostering a demand-side environment in which appropriate innovative competition can flourish. For example, permitting or increasing the flow of product information to consumers can encourage disease management programmes and foster patient empowerment.

To improve the operation of the demand side of the market, an understanding is required that pharmaceuticals as frequently purchased, low-priced items, are not the typical product traded in casualty insurance contracts. Insurance coverage is generally desirable when potentially large sums are at risk. It can even be mandatory (as with vehicle accident insurance) if there are third party effects (externalities). And it may be that some pharmaceutical products meet such criteria. Most do not. As a consequence pharmaceuticals should be bought and sold at full price by consumers acting, where appropriate, with the advice of their doctor. The benefits to the patient and to the economy would be large. In the UK Prescription charges (which already form a large proportion of a medicine's cost - £6 out of £10.55) could be abolished and with them a muscle-bound panoply of supporting controls. The administrative saving would permit

the NHS to focus its pharmaceutical budget on patients in need, not on products and prescriptions for all.

Government could then remove the Selected List of drugs which are reimbursable, and simplify the complex listing of currently exempt patient groups covering well over half the population. In short, abolition of the prescription charge implies that patients pay chemists directly for medicines prescribed. This in turn would encourage price competition for pharmaceuticals at retail level (in contrast to suppressing distribution channel innovations as has happened in the recent past). New distribution channels such as the dispensing of medicines for chronic conditions by mail order could emerge. Patients would select their distributor according to the package of price, service and product most convenient to them as individuals. Downward pressure would be exerted on pharmaceutical prices through the distribution channels back to the manufacturer. Brands or generics would be selected according to patient preference and doctor advice. Over-the-counter medicines would be selected instead of prescription drugs as appropriate. Self-medication, without the doctor's advice, would expand to the level deemed appropriate by patients themselves and not be critically constrained by the distribution of 'free' substitutes in the form of prescription medicines.

The PPRS would become redundant as the NHS ceased to be the industry's main purchaser and policy schizophrenia over whether the industry should be encouraged by industrial policy or somehow constrained to contain the health care budget would be removed. Anti-competitive actions by the industry would, of course, still remain subject to investigation by the Office of Fair Trading as is the entirety of British industry. They are consequently less likely selectively to distort the framework of incentives surrounding the industry relative to all other industries either on the supply side when it purchases labour, capital or material inputs, or on the demand side when it sells to its main domestic or overseas customers.

Increasing the power of the patient will inevitably put further pressure on distribution channels and blur the edges between what drugs ought to be available with or without a doctor's prescription. Some doctors may wish to enter dispensing on a larger scale than they do at present. Some pharmacists, on the other hand, may wish to assume more responsibility for advice provision and product recommendation. At the moment (except for certain categories of product) pharmacists may not discuss symptoms, dispense and so recommend product purchase. There is no obvious reason, in an increasingly sophisticated environment where patients are increasingly anxious to assume responsibility for their welfare (and the state is increasingly anxious to withdraw from its earlier self-imposed financial commitments), for not questioning these boundaries.

But what of the chronically sick where a casualty insurance argument can be made? Green and Lucas argue quite simply for exemption from charges for the poor, and catastrophic insurance for the remainder of the population. Thus existing exempt groups such as the elderly middle class, the children of the well-off, the diabetic who suffers an acute illness unrelated to his diabetes, and others would cease to be exempt. A market would therefore develop with most people most of the time paying market-related prices, while the indigent would continue to receive 'free' or 'near-free' medicines and those with larger outlays would have stop-loss insurance cover.

The additional advantage of separating out social insurance from casualty insurance is that more is available for redistribution. The poor can be made incrementally and relatively better-off than under the current system.

Since innovative rivalry is in pharmaceutical markets, its price-depressing competitive influence should be allowed its full effect and not distorted or suppressed by regulation at the level of either the pharmaceutical manufacturer or the retailer. Rather, patients and prescribers should be empowered to practice cost-conscious consumption by meaningfully exercising cost-effective demand for medicines. This can be facilitated by information provision, by a greater reliance on price and by encouraging, not discouraging, sound disease management principles at insurer, prescriber and patient levels. While in countries like the US and South Africa the tax break for employees' insurance premiums, which distorts demand, and which benefits richer rather than poorer patients, should be abolished.

