

Differential Pricing for Pharmaceuticals: Reconciling Access, R&D and Patents

Patricia M. Danzon PhD.
The Wharton School
University of Pennsylvania

December 2001

Table of Contents

1. Introduction and Summary	1
2. The Cost of Research-Based Pharmaceuticals and the Role of Patents.....	5
The Economic Role of Patents.....	6
3. Efficient Payment for R&D: Ramsey Pricing.....	7
Ramsey Price Differentials vs. Profit-Maximizing Differentials.....	9
4. Actual vs. Optimal Price Differentials and the Breakdown of Separate Markets: Parallel Trade and External Referencing.....	10
Actual vs. Optimal Price Differentials.....	10
The Breakdown of Separate Markets: Parallel Trade and External Referencing..	11
Differential Pricing Does Not Imply Cost-Shifting.....	12
5. Policies to Maintain Separate Markets and Price Differentials.....	13
Defining Patents Based on National Boundaries, Including the Right to Bar Parallel Trade.....	14
Implementing Differential Pricing through Country-Specific Contracts with Confidential Rebates.....	14
High Income Countries Should Forego Regulation Based on Foreign Prices.....	16
6. Compulsory Licensing.....	18
7. Conclusions.....	19
8. References.....	20

1. Introduction and Summary

Developing countries have two primary needs in access to medicines. The first is access to medicines that target diseases that are prevalent in both high and low income countries (hereafter “HI drugs”) at prices developing countries can afford, with distribution systems and health care infrastructure to assure effective use. The second need is for the development of new medicines to treat diseases that exist only in less developed countries (“LDC drugs”). The standard means to encourage innovation in any industry, including pharmaceuticals, is to provide patent rights. But patents work by enabling originator firms to charge prices above marginal cost, in order to recoup their investments in R&D. Thus prima facie it appears that patents inevitably lead to high prices and that there is an inherent conflict between the objective of encouraging R&D and the objective of assuring access at affordable prices.

This paper argues that differential pricing is the key to resolving the potential conflict between patents, which are necessary to preserve incentives for R&D, and affordability of drugs in developing countries. Differential pricing is defined here to refer to a specific pricing structure and implementation strategy. Specifically, differential pricing as proposed here involves permitting manufacturers to charge different prices in different markets, based on “Ramsey pricing” principles, which imply prices inversely related to price sensitivity or demand elasticity.¹ Implementation requires the adoption of policies to prevent low prices in low-income countries from “leaking” from the low-income countries to higher income countries, thereby undermining the potentially higher prices in these high income countries. Such leakages occur primarily due to parallel trade or “external referencing.”² Differential pricing as proposed here does not require regulation by some supranational body to set specific discounts for specific drugs in specific countries. Such regulation is likely to be counterproductive to obtaining appropriate price differentials.

The economic foundations of differential pricing, as laid out here, are not a defense of existing price differentials – on the contrary. Current prices for some on-patent drugs may appear to be inappropriately high, relative to income, in some low income countries. However, this is because current conditions do not encourage appropriate price differentials. Thus the status quo is not a basis for judging the potential for differential pricing combined with patents. Given the extent of current cross-market leakages, manufacturers would rationally be reluctant to grant low prices in low-income countries. Thus current prices and price differentials are not a reliable indicator of how differential pricing might work, if

¹ The price elasticity of demand is defined formally as the percentage change in quantity in response to a percentage change in price.

the necessary implementation steps were taken. In fact, differential pricing could be an efficient and equitable way to reconcile the two seemingly conflicting objectives, of making drugs affordable for developing countries, while preserving patent protection and incentives for R&D.

Before turning to the analysis of appropriate differential pricing, it is important to clarify what it can and cannot do. Sound medicines policy for less developed countries (LDCs) must distinguish between HI drugs, with demand in high income countries, and LDC drugs, that target diseases that exist only in less developed countries. Differential pricing is most relevant for HI drugs.

For HI drugs with an affluent country market, differential pricing can simultaneously achieve affordability in LDCs and preservation of incentives for R&D. The key is that, under well-designed differential pricing, prices in affluent (and, to a lesser extent middle income countries) exceed the marginal cost of production and distribution in those countries by enough, in aggregate, to cover the joint costs of R&D, while prices in LDCs cover only their marginal cost. Antibiotics and HIV-AIDS drugs exemplify medicines that serve both HI and LDC markets, for which differential pricing could simultaneously yield prices that are affordable to low income countries while preserving incentives for R&D.³

For drugs to treat diseases found only LDCs, there is no HI market where prices can exceed marginal costs in order to cover the joint costs of R&D. LDC drugs are, by definition, drugs for which the prices that can be afforded by LDC patients are insufficient to cover costs and yield an incentive for innovators to invest in R&D. Thus some external subsidy— either a demand-side subsidy to patients or a supply-side subsidy to innovator firms – is necessary to create incentives to develop treatments for LDC-only diseases. Patents are useful but will not suffice: having the legal authority to charge high prices is of no value if patients cannot pay. Various subsidy options have been proposed for funding R&D on LDC drugs, but these are not discussed here. The focus of this paper is on the use of differential pricing for drugs that serve both HI and LDC markets.

The form of differential pricing proposed here is based on the economic literature on “Ramsey pricing” (Ramsey, 1927), which was developed to address the problem of paying for joint costs that simultaneously serve many consumers. Pharmaceutical R&D is such a joint cost. Ramsey pricing implies prices that vary across markets inversely with each market’s price sensitivity or demand elasticity. Although price elasticity is not directly observable, per capita income is an important determinant and

² “External referencing” refers to the practice whereby a government in country A refers to (usually lower) prices in country B as a benchmark for regulating prices in country A. Such external referencing can be formally built in to the regulatory system, as in Canada, Italy and the Netherlands, or can be an informal benchmark.

³ Even with prices at marginal cost in LDCs, the neediest patients may require additional subsidies for chronic medicines and those with high production costs. In these cases differential pricing can still be an important part but not the whole solution.

can serve as a good approximation, possibly adjusted by burden of specific diseases such as HIV-Aids. Thus this approach would yield low prices in low-income countries, particularly those with a high disease burden, and higher prices in high-income countries.

A necessary condition for maintaining appropriate price differentials is that markets are separable, such that if manufacturers charge low prices in LDCs, these low prices do not spill over to HI markets. In practice, markets are not separate because governments in many countries regulate their domestic prices based on lower prices in other countries (“external referencing”) and intermediaries can import products from low to higher price countries (“parallel trade”). Given these linkages across markets, basic economics predicts that manufacturers will rationally seek to maintain much higher prices in LDCs than they would require if markets were separate and price leakages did not occur. Thus as long as low prices offered to LDCs can spill over to other, potentially higher-price markets, manufacturers will rationally be reluctant to offer prices at marginal cost in LDCs, except when forced by political or other pressure. Price leakages, rather than patents per se, plausibly explain why prices for patented drugs often appear to be set at unaffordable levels in low-income countries.

A major conclusion of this analysis is that to assure low prices in LDCs requires that higher-income countries abstain from trying to “import” low LDC prices and that policies be established that enforce such market separation. Given enforceable barriers between markets, as discussed below, it will plausibly be in the self-interest of pharmaceutical manufacturers to charge prices close to marginal cost in LDCs, which would go a long way towards making drugs more affordable. Differential pricing does not require a supranational regulatory system to establish and enforce price differentials, and such a system is not recommended here. Rather, a system of confidential rebates, negotiated between purchasers (governments or other third party payers) and individual companies, is more flexible and more likely to simultaneously achieve appropriate price differences and preempt parallel trade and external referencing. By contrast, if differentials are regulated by a supranational body there is nothing to prevent these practices that are fatal to appropriate differential pricing.

Developing the institutional framework necessary to support differential pricing is particularly important now as LDCs move to adopt the TRIPs (Trade-Related aspects of Intellectual Property) patent provisions of the World Trade Organization (WTO). Such patents would generally bar cheap copy products for the life of the patent. Patents are compatible with affordable prices of on-patent, originator drugs in LDCs, provided that the low prices in LDCs do not spillover to HI markets. Patent provisions are likely to permit compulsory licensing and/or importing of copy products in cases of “national emergency” and this may be appropriate provided that “national emergency” is narrowly defined. But there is a risk that “national emergency” will be defined broadly to apply to most on-patent medicines,

thereby nullifying the purpose of patent protection, unless conditions are established such that differential pricing can work to make on-patent drugs affordable in low income countries. This requires adopting policies to prevent price leakages. If originator firms can be confident that low prices will stay in low income countries, then their prices for originator products may be no higher than prices on copy products, assuming similar marginal costs.

The case for compulsory licensing of local generic manufacturers, to produce copy versions of on-patent products, assumes that these generic products would be significantly cheaper than the patented originator products. However, it is possible that the main reasons that compulsory-licensed products are cheaper are either lower quality or that the manufacturers of compulsory-licensed products are not concerned about price spillovers to higher-price markets because they cannot legally sell there. If so, then a potentially more efficient alternative to compulsory licensing would be to prevent the leakage of low prices from LDC markets, such that originator firms would have incentives to price low in LDCs, without sacrificing the higher prices in HI markets needed to pay for R&D.

This proposal for differential pricing and a bar on parallel trade may seem at odds with standard economic principles. Simple economic theory concludes that there are efficiency gains from free trade and competition, and that the efficient outcome is a single price, equal to marginal cost, in all markets. However, this standard competitive model is not appropriate for research-based pharmaceuticals or any other industry with high fixed, joint costs of R&D that cannot be sustained under marginal cost pricing. In this case, the more relevant model is the Ramsey pricing model, which concludes that differential pricing is the most efficient approach to paying for pharmaceutical R&D that serves consumers with widely differing ability to pay.

The economic case for differential pricing of pharmaceuticals is thus based not only on equity but also on considerations of economic efficiency (maximizing the utility produced from available resources). This efficiency case for differential pricing is widely misunderstood. Consumers in higher income countries of Europe and North America may appear to benefit in the short run from trying to “import” the lower prices available in low-income countries, through either parallel trade or other regulation based on lower foreign prices. But in the long run high-income countries will be worse off under a uniform pricing regime than under differential pricing, in the likely event that there is less revenue, hence less R&D and fewer new drugs under uniform pricing than under differential pricing. Thus the efficiency conclusion is that everyone can be better off under differential pricing if more drugs are developed and those that are developed are more widely used in low-income countries.⁴ Differential

⁴ This efficiency conclusion takes existing patent and insurance regimes as given and appropriate, and hence assumes that any product that can generate revenue sufficient to cover all costs, including R&D, yields benefits greater than

pricing is generally also consistent with conventional notions of equity, because it implies lower prices in low-income countries than in higher income countries. By contrast, the current environment creates pressure towards uniform pricing, which results in inappropriately high prices for on-patent drugs in low-income countries, or delay or failure to launch the drugs if countries cannot pay these higher prices, even though they could pay prices that cover their marginal cost.

The structure of the paper is as follows: Section 2 reviews the importance of joint costs in the cost structure of the research-based pharmaceutical industry. Section 3 outlines the theory of Ramsey pricing, and compares these Ramsey-optimal price differentials to the price differentials that in theory emerge in monopolistically competitive markets with entry. Section 4 examines the determinants of actual price differences within the US and cross-nationally, and reviews the effects of parallel trade and the cost shifting argument against differential pricing. Section 5 discusses implementation of differential pricing and Section 6 discusses compulsory licensing. Section 7 concludes.

2. The Cost Structure of Research-Based Pharmaceuticals and the Role of Patents

The research-based pharmaceutical industry in the US spends 13-20 percent of sales on R&D, compared to under 4 percent for US industry overall (PhRMA, 2001). This sales-based measure understates R&D expense as a percentage of the total costs of developing and producing new drugs, because it omits the “opportunity” or capital cost of funds over the 8-12 years required for drug discovery and development.⁵ Adding in this cost of funds, R&D accounts for roughly 30 percent of the total cost of developing, producing and marketing new drugs, with all costs measured as discounted present value at the time of product launch (Danzon, 1997).

This large R&D expense complicates pricing for several reasons. First, R&D is a fixed, globally joint cost; that is, this cost is largely invariant to the number of patients or countries that ultimately use the drug and cannot be causally attributed to specific countries. Once R&D has developed a compound to serve affluent countries, no incremental R&D expense is needed to serve low-income countries.⁶ Second, this global joint cost is largely sunk by the time the product is launched and price is negotiated. Marginal cost (MC), which is the incremental cost incurred to serve an additional country, patient group or individual patient, depending on the decision at hand. As a drug advances through its life cycle and is

costs and so should be produced. If it is believed that existing patent or insurance regimes encourage too much R&D, these should be addressed directly, not indirectly through restrictions on differential pricing.

⁵ The opportunity cost is the highest alternative return that the company could have realized on the funds invested. See DiMasi et al., 1991.

launched in more countries, country-specific launch costs are sunk and marginal cost includes only the incremental costs of production and sales of additional units, which are a small fraction of total costs.

If there were no patents, generically equivalent copy products could enter freely and competition would force prices down to marginal cost. Marginal cost pricing would suffice to cover the expenses of copy products that incur only production and distribution costs with negligible R&D or promotion expense. But marginal cost pricing cannot generate revenue sufficient to cover the R&D costs of innovator firms. Hence free entry and the resulting marginal cost pricing are incompatible with sustained incentives for R&D.

The economic role of patents

The economic purpose of patents is to bar entry of copy products for the term of the patent, to provide the innovator firm with an opportunity to price above marginal cost and thereby recoup R&D expense, in order to preserve incentives for future R&D. Strict economic theory views patent protection as a “second best” way to pay for R&D. In a “first best” or fully efficient outcome, all consumers whose marginal benefit exceeds marginal cost should use the product; however, patents permit pricing above marginal cost, hence some consumers may forego the product even though their marginal benefit exceeds the marginal cost. But with large fixed costs of R&D no first best solution is possible: the first best requires marginal cost pricing to consumers, but this would generate inadequate revenue to sustain innovation unless the government subsidizes R&D. However, allocating subsidies in a way that creates efficient incentives and avoids waste is difficult, if not impossible, and raising the necessary taxes undermines efficiency and possibly equity in other sectors of the economy. Thus a patent system, which enables innovator firms to charge prices above marginal cost to consumers who use the product, is generally viewed as the best practical approach to funding R&D. Moreover, in the case of patents for health care products, health insurance in most industrialized countries reduces the out-of-pocket price to the patient far below the full price of a drug, which greatly reduces the likelihood that pricing above marginal cost will lead to suboptimal consumption.

In LDCs that lack extensive health insurance, there is a concern that patents lead to prices significantly above marginal cost and hence to suboptimal utilization. However, this is not an inevitable result of patents. Patents may make pricing above MC theoretically possible, but this may not be in the firm’s self-interest if consumers cannot afford to pay. Leaving aside altruistic motivations, rational pricing strategy for the firm depends on demand elasticity, which depends on many factors including income and availability of substitute products. If, as casual evidence suggests, low income in LDCs

⁶ Drug discovery is a pure joint cost. Drug development, including clinical trials to prove safety and efficacy, is increasingly a joint cost with the harmonization of requirements and conduct of multi-country trials that are used for

makes demand highly elastic, such that volume is very responsive to price, then a patent-holder would rationally set prices near marginal cost in the absence of leakages to other, less-elastic markets. Demand will tend to be more elastic in therapeutic classes with multiple competing products, which is the norm for most pharmaceuticals. But even drugs with few competitors may face elastic demand in low income countries, particularly if governments/purchasers bargain on behalf of their populations to offer higher volume in return for lower prices, thereby making demand elasticity explicit. More on this below.

3. Efficient Payment for R&D: Ramsey Pricing

Necessary conditions for (second best) efficiency in drug utilization and drug development are: (1) price P is at least equal to MC in each market or country; and (2) prices exceed MC by enough, in aggregate over all markets, to cover the joint costs of R&D, including a normal, risk-adjusted rate of return on capital (F):

$$P_j \geq MC_j, \text{ and} \quad (1)$$

$$\Sigma (P_j - MC_j) \geq F \quad (2)$$

The first condition, that price covers marginal cost in each market, assures that marginal benefit exceeds marginal cost, as required for efficient resource allocation. In the case of health services, the price paid for drugs may include social insurance and possibly other subsidy payments, reflecting the willingness of higher income taxpayers or countries to subsidize consumption for more needy patients. The second equation is both a break-even condition for the firm and a necessary condition for the efficient allocation of resources to R&D. These necessary conditions for efficiency in drug consumption and innovation do not imply or require that prices should be the same for all consumers.

An important remaining policy question is, What pricing structure across markets would satisfy these two conditions and yield the greatest social welfare for consumers? Ramsey pricing (Ramsey, 1927; Baumol and Bradford, 1970) provides the answer to this question. Ramsey pricing theory was developed to address the problem of optimal pricing of utilities, which incur large joint capacity costs plus low consumer-specific marginal costs. The utility capacity is fundamentally similar to the joint costs of pharmaceutical R&D, hence the same theory applies.⁷

regulatory submissions in many countries.

⁷ Although Ramsey pricing in the utility context has often been implemented by regulation, this is not essential. The optimality of the Ramsey pricing structure is separate from the issue of how it is implemented. As discussed further below, utilities were traditionally local monopolies serving local customer groups -- conditions in which regulation of the price differentials is most likely to work reasonably well. Regulation of price differentials is less necessary and less likely to work well in the case of pharmaceuticals, which are marketed in many different countries and typically face competition from substitute products.

Ramsey optimal price (ROP) differentials are designed to yield the highest possible social welfare, subject to the condition that the producer achieves some target profit level, usually a normal, risk-adjusted return on capital. The Ramsey pricing solution is that prices should differ across customer groups in inverse relation to their demand elasticities and on average exceed short run marginal cost, in order to cover the joint fixed cost. In the case of a single product, the condition for the optimal price markup for submarket j is:⁸

$$\frac{p^j - c^j}{p^j} = - \frac{\lambda}{1 + \lambda} \frac{1}{E_j} \quad (3)$$

or $L^j = D / E_j \quad (3)'$

where E_j is the own elasticity of demand in market j . Thus L^j , which is the mark-up of price over marginal cost (also called the Lerner index) in market j , should be proportional to the demand elasticity E_j , with the proportionality term D defined by the breakeven constraint. Thus if marginal cost is the same in all markets, Ramsey optimal prices differ depending only on demand elasticities. If marginal cost differs across markets, these conditions apply to mark-ups over market-specific marginal cost.

The intuitive explanation for Ramsey pricing is simple. Recall that the ideal would be to charge everyone their marginal cost but this is not practical because pricing at marginal cost would not cover R&D. The Ramsey solution minimizes the loss from departing from this ideal: more price-sensitive users should be charged a smaller mark-up over marginal cost than less price insensitive users, because the price-sensitive users would reduce their consumption by proportionately more, if faced with the same prices. Charging lower prices to more price-sensitive users is also consistent with equity, assuming that price-elastic demand results from low income.

Ramsey Price Differentials vs. Profit-Maximizing Differentials

One common objection to Ramsey pricing is that it proposes price differentials similar to those charged by a price discriminating monopolist (PDM). The monopolist's profit-maximizing mark-up in market j is:

$$(p^j - c^j) / p^j = L^j = 1/E_j \quad (4)$$

⁸ With multiple products and nonzero cross-price elasticities, optimal price mark-ups should take into account these cross-elasticity effects; with multiple firms, strategic interactions by firms should also be taken into account (Breutigam, 1984; Laffont and Tirole, 1993; Prieger, 1996; Danzon, 1997).

Comparing the price markups in eqs. (3) and (4), the relative markups across markets are the same under PDM as under ROP, but the absolute prices may differ due to the profit constraint factor, D . Ramsey prices are derived to yield a specific target return on capital for the firm -- usually a normal, risk-adjusted rate of return. By contrast, the unconstrained monopolist may try to maximize profit, but may actually realize more or less than a normal rate of return in any given year. But in the long run, with unrestricted entry and exit of firms offering competing but differentiated products, dynamic competition will reduce expected profits to normal levels at the margin. This is simply the standard monopolistic competition model, which fits the pharmaceutical industry reasonably well. Entry occurs such that excess expected profits are eliminated for the marginal firm and the marginal product in each firm's portfolio of products. Ex post, actual realized profits of a given firm may be above or below normal levels. Given the scientific and market risks faced by the pharmaceutical industry, it is not surprising that expectations in pharmaceuticals are not always accurate. Many products yield less than a normal return on the capital invested and others yield more (Grabowski and Vernon, 1990); some firms have been very successful while others have been merged out of existence, and average profitability has varied across time and countries.

In fact, the similarity between the welfare maximizing (Ramsey optimal) pricing structure and price discrimination pricing structure is not surprising and is fortuitous. It means that firms, pursuing their own self-interest, will be led to set price differentials across markets that are socially appropriate. In theory, price discrimination is efficient, since all consumers with marginal value greater than marginal cost consume the product. There may be an objection on grounds of equity, if these prices yield excess expected profits to the monopolist, after adjusting for risk. But in an industry such as pharmaceuticals, with continuous entry of new firms and products that are close substitutes, any excess expected profit will be bid away. Under these conditions, the market structure is such that, if markets are separate, firms would set price differentials that are inversely related to demand elasticities. These differentials would be in the self-interest of the firm but are also potentially Ramsey-optimal and can be constrained by competition and price-conscious buyers to yield normal expected profits.

4. Actual vs. Optimal Price Differentials and the Breakdown Of Separate Markets: Parallel Trade and External Referencing

Actual vs. Optimal Price Differentials

Opposition to the differential pricing approach is based in part on the observation that actual price differences do not appear to approximate appropriate levels, given income differentials. In the US some low-income individuals face relatively high prices, while internationally some low-income countries

seem to face prices that are unaffordable and significantly above marginal cost, in the absence of political pressure. In fact, these observations show that the current system is not well designed to achieve appropriate price differentials; they do not show how the approach might work if the necessary reforms were adopted.

In the US, actual price differentials between market segments for on-patent drugs are reasonably consistent with inverse demand elasticities. Health plans that manage their pharmacy benefits, using tightly managed formularies with lists of preferred drugs, effectively increase the demand elasticity facing sellers and hence are able to negotiate significant discounts for branded products. By contrast, patients who have unmanaged plans or no insurance do not get the same discounts. The reason is that they have no price-sensitive intermediary that can shift market share towards firms that offer lower prices.⁹ Although in theory physicians might play this role, in practice physicians' prescribing decisions appear to be relatively price-insensitive. These are fundamentally insurance problems that are best addressed by insurance reforms, not by regulating drug prices. Insurance reforms that extend to seniors and other low income individuals a choice of managed drug plans would enable them to benefit from negotiated discounts on drug prices and dispensing fees.

Competitive discounting in the US has been constrained since 1991 by the Medicaid "best price" provision, which requires manufacturers of branded products to give the public Medicaid program the largest discount that they give to any private customer. But Medicaid demand is relatively price-inelastic: it traditionally does not shift market share to products that give lower prices, unlike managed private plans. Thus the effect of linking Medicaid's relatively price-inelastic market to the more price-elastic private market has been to reduce discounts that manufacturers are willing to grant to private buyers.¹⁰ Essentially, the Medicaid best price provisions result in leakage of low prices from more price-elastic to less price-elastic markets. Thus in the US as in the international context, leakages from more elastic to less elastic markets tend to erode discounts in the more price-elastic markets.

Turning from the US to cross-national price differentials, these appear to deviate significantly from what might be expected based on income as a proxy for price sensitivity: some high income countries have relatively low prices, while some low-income countries face high prices relative to their income level. Several factors contribute to this. First, regulators in some high-income countries use their

⁹ This is explained in more detail in Danzon (1997).

¹⁰ Formally, given the Medicaid best price provision (or linkage between any two markets), the firm will set the price based on a weighted average of the elasticities in the two separate markets. If the less elastic market is significantly larger, this will dominate the common price and the more elastic market will face a higher price than it would if markets were separate.

bargaining leverage -- sometimes combined with external referencing -- to reduce their prices to relatively low levels, leaving others to pay for the joint costs of R&D.

Second, the tendency for prices in low-income countries to be inappropriately high, relative to their average per capita income, probably reflects primarily manufacturers' concerns over price leakages -- that granting lower prices to low income countries would undermine potentially higher prices in other countries. In addition, the highly unequal distribution of income in some countries may mean that a small, high-income subgroup dominates potential pharmaceutical sales, leading to prices that are geared to that subgroup but are unaffordable for other subgroups.¹¹ The appropriate solution in such cases is to separate the submarkets within the country, for example, by establishing a program that serves the low-income subgroup only, without spillovers to the higher income subgroup, such that prices could differ between subgroups within the country, as well as cross-nationally.

The Breakdown of Separate Markets: Parallel Trade and External Referencing

The breakdown of separate markets and hence of manufacturers' ability to maintain price differentials is probably the single most important obstacle to lower prices in low-income countries. The primary factors are two policies that are favored by higher-income countries: parallel trade and external referencing. Parallel trade occurs when an intermediary exports an originator product from one country to another to profit from the price differentials set by the manufacturer. Parallel trade violates traditional patent rules, whereby the patent holder could bar unauthorized importation of its product. These traditional patent rules were preserved in the North American Free Trade Association (NAFTA). However, the European Union authorizes parallel trade within the EU, adopting the view that the originator firm exhausts its patent rights once it places the product on the market anywhere in the EU. Parallel trade has become less costly with the establishment of the European Medicines Agency in 1995, which harmonized regulatory approval requirements, including labeling and packaging for all EU countries, and with adoption of the euro which makes price differences more transparent. The US recently enacted provisions to permit re-importation of drugs. This legislation has so far not been implemented, due to concern over assuring quality of imports and doubt about whether cost savings would be passed on to consumers or simply accrue to middlemen -- a problem faced by all parallel importing countries. However, this and other broader plans for parallel trade are being considered, both in the US and other countries.

Parallel trade is often erroneously defended using the standard economic arguments for free trade, but in fact the standard gains from trade do not apply in the case of parallel trade. Free trade usually yields efficiency gains because the exporting country has lower production costs, due to either superior

¹¹ I am indebted to Jaishree Watal for emphasizing this point.

techniques or lower input costs. But lower prices in countries that parallel export pharmaceuticals usually result from either aggressive price regulation, lack of patent protection, or lower per capita income which leads the originator firm to grant lower prices.¹² None of these factors creates an efficiency gain from trade. In fact, parallel trade can increase social costs, due to costs of transportation, relabelling and quality control. Moreover, at least some of the savings usually accrue to the middlemen, not to the consumers or payers in the importing country who continue to pay the higher price.¹³

The second policy that erodes separate markets and promotes price spillovers is external referencing, which occurs when governments or other purchasers use low foreign drug prices as a benchmark for regulating their domestic prices. Such external referencing is used formally by the Netherlands, Canada, Greece and Italy, among others, and used informally by many other countries.¹⁴ External referencing is equivalent to fully importing a foreign price. The risk that low prices granted in low-income countries would lead high-income countries to demand similarly low prices is probably the single most important obstacle to lower prices in these low income countries.

Faced with price leakages due to external referencing and parallel trade, a firm's rational response is to attempt to set a single price or narrow band of prices. Consistent with this prediction, companies frequently now attempt to obtain a uniform launch price throughout the EU, and launch may be delayed in countries that do not meet this target price. Formally, if two markets L and H are linked, the profit-maximizing strategy is to charge a single price P in both markets, where P is based on the weighted average of the elasticities in the two markets, with weights that reflect relative shares of total volume Q:

$$(P - MC)/P = 1/(E_h w_h + E_l w_l) \quad (5)$$

$$\text{where } w_l = q_l / Q \text{ and } w_h = q_h / Q$$

Thus if the low income market is small, relative to the high income market, the single price will be dominated by conditions in the high income market and hence will appear unaffordable in the low income country. This single price will also far exceed the price that would have been charged, had markets been separate, as determined by eq. (4).

This breakdown of price differentials that are appropriate to the different conditions in each market is inefficient and inequitable. Consumers in low-income countries face inappropriately high

¹² Lower labor cost is only a small fraction of total production costs, hence is unlikely to account for significant price differences.

¹³ The UK and the Netherlands attempt to "claw back" the profit that accrues to the pharmacy when it dispenses a cheaper parallel import rather than brand.

¹⁴ President Clinton's Health Security Act (1994) proposed to limit US prices to the lowest price in 22 countries. If enacted, this would have permitted low prices in New Zealand, for example, to erode revenues in the US.

prices and forego medicines, even though they might be willing to pay prices sufficient to cover their marginal cost. High-income countries might appear to benefit in the short run from trying to import low prices. But in the long run these countries are also likely to lose as the break-down of differential pricing leads to lower revenues, less R&D and hence fewer new medicines.

Differential pricing does not imply cost-shifting

A common objection to differential pricing is that it implies “cost shifting” from low-price to high-price markets. This argument either ignores the jointness of costs or mistakenly assumes that joint costs should be allocated equally to all users. It also assumes irrational behavior on the part of manufacturers. As long as markets are separate, a firm would rationally set the price in each market based on conditions in that market, independent of prices in other markets. To do otherwise would actually reduce net revenues. Moreover, for the long run decision of deciding whether to develop a new drug, the simple breakeven rule compares expected total revenues to total costs. If low price users cover at least their marginal costs and make some even small contribution to the joint costs of R&D, prices in high price countries can actually be lower than they would have to be to cover joint costs in the absence of contributions from the low price countries. Thus for costs that truly jointly benefit different consumer groups but must be incurred to serve any one of them – which is the case for pharmaceutical R&D -- the terminology of cost shifting is misplaced. Moreover, the behavior ascribed to manufacturers, of raising prices in one market because of lower prices in another market, would be contrary to their self-interest, as long as markets are separate.

It is true that if price differences are unsustainable, due to parallel trade and external referencing, then manufacturers will tend to charge a single price that is between the prices that would have been offered, had markets been separate. Under such uniform pricing, consumers with relatively inelastic demand may have somewhat lower prices due to associating with consumers with more elastic demand. Although the high-income, inelastic users may try to justify this as “eliminating cost-shifting,” it could more appropriately be called “free riding” by the high-income, price-inelastic consumers on the low-income, price-elastic consumers.

5. Policies to Maintain Separate Markets and Price Differentials

This analysis concludes that policies that promote drug prices that differ across markets based on differences in true price elasticity will lead to higher overall social welfare and greater equity than policies that lead to narrow pricing bands or uniform prices across countries. Moreover, firms would tend to adopt roughly appropriate price differentials, even without supranational regulation, provided that markets are separate, such that low prices offered to low-income countries do not spillover to higher

income countries. Markets are currently not separate, because higher and middle income countries use parallel trade and external referencing to try to import lower prices, which in turn makes companies reluctant to grant really low prices to low income countries. Differential pricing, including appropriately low prices in low-income countries, will only be possible if higher income countries accept the responsibility to pay higher prices, foregoing the temptation to try to obtain the lower prices granted to low income countries. Specific policies that could help sustain price differentials include:

(a) Defining patents based on national boundaries, including the right to bar parallel trade

The simplest way to stop parallel trade is to define patents to include the right for a patent holder in each country to bar unauthorized imports of products that are under patent protection. This is consistent with traditional patent rights in the US and in Europe with respect to non-member states. The economic efficiency case for national boundaries for patents (no doctrine of international exhaustion) is strongest for industries such as the pharmaceutical industry, which incurs significant global, joint R&D expense that is optimally recouped by differential pricing. The fact that pharmaceuticals are often subject to price regulation further strengthens the case for permitting patent holders to bar parallel trade across national boundaries, since price regulation increases the likelihood that countries have low prices due to aggressive regulation rather than lower real production costs. Note, however, that national patents would prevent parallel trade but would not stop spillovers due to external referencing.

(b) Implementing Differential Pricing through Country-Specific Contracts with Confidential Rebates

External referencing, which is the more serious source of price spillovers, can best be addressed by making low prices granted to one purchaser unobservable to others, such that they cannot be copied. This can be done if manufacturers grant lower prices in the form of confidential rebates, to be paid by the manufacturer directly to the purchaser/payer, rather than by selling to distributor/wholesalers at the lower price. If discounts to low income countries or market segments are given in the form of confidential rebates paid directly to the ultimate purchaser, while wholesalers are supplied at a common price, this eliminates the opportunity for other purchasers to demand similar rebates. In addition, it eliminates the opportunity for wholesalers or other intermediaries to purchase product at the low price intended for low-income countries and export it to higher-price countries. Thus implementing price differentials through confidential rebates to final purchasers prevents price leakages due to both parallel trade and external referencing and confines the discounts to the intended beneficiaries. Moreover, if these discounts are part of a negotiated contract which links specific discounts to specific volume of use, this would enable discounts to be targeted to low income countries or other users who demonstrate price elasticity through

their willingness to accept volume-price contracts. By making rebates payable *ex post*, depending on volume of use, difficulties of determining elasticities *ex ante*, due to bluffing and other bargaining strategies, are eliminated.

Such confidential discounts, often payable *ex post* on demonstrated ability to shift market share, are the chief means by which managed care purchasers in the US get lower prices. Discounts can be targeted to those payers with truly elastic demand and other, less-elastic purchasers cannot demand similar discounts because the discounts are not known. The only exception is that the federal government uses its legislative authority to demand best price discounts for Medicaid and other federal programs. However, if confidential discounts were used to achieve price differences across countries, no country could insist on knowing discounts given to other countries, so price leakages would be prevented. Targeted rebates were also used to maintain lower *ex post* prices in East Germany than West Germany after reunification. Targeted rebates, with discounts linked to volume of product use, are also common in other industries.

It may be objected that determining appropriate rebates based on true demand elasticities is problematic, since purchasers who are third party payers may not know the true preferences of their consumers/enrollees and may have incentives to conceal their true preferences to bargain for a lower price. However, this is no different from any other negotiated contract, where seller and buyer negotiate to reach a mutually acceptable outcome, given their respective interests. Making rebates payable *ex post*, contingent on actual volume of use, circumvents some of these problems.

Admittedly negotiating and enforcing contracts entails administrative expense and may yield price differentials that are only approximately Ramsey-optimal. However, there is no perfect and costless solution to the problem of simultaneously paying for R&D while making drugs affordable to low-income countries. Compared to the status quo, with drug prices within a narrow band for a range of countries with very differing abilities to pay, the proposed policy of permitting confidential discounts would achieve lower prices for low income/price-sensitive countries. Moreover, everyone would benefit in the long run from a greater supply of innovative new pharmaceuticals.

The proposed system of price differentials through confidential rebates will only work if these discounts are negotiated between individual companies and purchasers. If discounts are regulated by a supranational body or regulatory authority, for a specified list of products in specific countries, such discounts are no longer confidential. If country A's discounted price is public information, then country B can use it for external referencing, as currently occurs, which will erode companies' willingness to grant discounts to low income countries. Moreover, if discounts are set uniformly for all products, this eliminates the purchaser's ability to use competition between companies to negotiate a larger discount

from one company in return for increasing its market share. Companies are likely to collectively resist a regulated schedule of discounts, even though individually they might grant larger discounts on specific products to individual countries that demonstrate ability to increase their market share. Thus discounts that are negotiated individually between companies and purchasers are likely to result in greater flexibility, more competition, lower prices in low income countries and less political wrangling than an attempt to define a regulated schedule of discounts that applies to specified products for specified low-income countries, with no means of preventing other, middle and high-income countries from demanding similar discounts.

A common objection to this approach is that small countries would have no bargaining power so would face high prices. But if the small country truly has elastic demand, then it is in the seller's self-interest to charge a price close to marginal cost, provided that this price remains in the small country. Small size only becomes a problem when the small country's price could spillover to a larger country that has less elastic demand. Given such leakages, the firm's best single price will be dominated by the large country's less elastic demand, as shown in eq. (5). Small price-elastic countries or markets that are truly separate from other less-elastic markets should be able to negotiate very low prices, unless there are significant fixed costs of operating in these countries.

(c) High income countries should forego regulation based on foreign prices

Any institutional framework to preserve differential pricing will only work if higher income countries forego the temptation to try to reduce their prices by referencing lower prices in low-income countries. High-income countries currently use various approaches for regulating prices and a full evaluation of alternative approaches is not possible here. However, external referencing has the particularly unfortunate side-effect of making manufacturers unwilling to offer low prices to low-income countries.

Determining appropriate prices in practice is not easy, because ideal Ramsey prices depend on unobservable preferences of consumers and taxpayers, which may be imperfectly reflected through the intermediation of public and private payers and physician prescribers. In practice, pricing is probably best resolved through confidential negotiations between purchasers and sellers, as in other industries and markets, with the recognition in the case of high income countries that in the long run they must pay for R&D if it is to continue. Tools of cost-benefit and cost-effectiveness analysis can help estimate the value of new drugs, taking into account alternative therapies, and price-volume contracts, which make discounts contingent on increased market share, can avoid the need to agree *ex ante* on a specific price and can be used to promote competition between therapeutic substitutes.

Although some may be concerned that such negotiations between payers and companies would leave market power in the hands of companies, in fact the risk may be equally great that prices will be too low, due to short term horizons of purchaser/regulators who have monopsony power (are sole buyers) and may be more concerned with their current budgets than with assuring the future supply of medicines. On the seller side, true monopoly is rare and short-lived, despite patents, due to the rapid entry of follower products into most new therapeutic classes. Thus payers can exploit competition between alternative, similar drugs to limit prices to reasonable levels. Regulatory mechanisms that attempt to force prices down to short run marginal cost, if widely adopted, will be unsustainable in the long run.

A related argument is that, since Ramsey pricing was originally applied to the regulation of utilities, regulation is a necessary condition of implementing it for the pharmaceutical industry. However, while the pharmaceutical industry resembles utilities in have large joint costs and low marginal costs, which makes Ramsey pricing appropriate, these industries differs in other important ways. Utilities were usually local monopolies. But any market power enjoyed by individual drugs derives primarily from the intentional grant of patents in order to permit pricing above marginal cost. Competition from therapeutic substitutes makes pure monopoly rare and temporary. Competition can also be encouraged by the design of insurance arrangements, including incentives for consumers and physicians to be cost-consciousness. Thus the monopoly rationale for regulation does not apply.

It is worth noting that the global nature of pharmaceutical joint costs makes the pharmaceutical industry more vulnerable to aggressive price regulation than other regulated industries. Traditional utility pricing formulae generally explicitly recognized the need to provide a reasonable return on capital. Because the utility's production capacity was country-specific, local users could not free ride: if they did not pay for capacity costs, their future access to services would obviously be at risk.¹⁵ By contrast, the global nature of the joint costs of pharmaceutical R&D creates the incentive and opportunity for regulators in each country to free ride, paying only their marginal cost and leaving others to pay the joint costs. Moreover, the long lag between initiating R&D and bringing products to market means that even if current low prices do reduce R&D and hence the future supply of new drugs, it will be hard to attribute future lack of innovation to specific current policies or politicians.¹⁶

¹⁵ As these utilities expand across national boundaries, allocating joint costs across countries may become more problematic, and problems may arise similar to those already experienced by pharmaceuticals.

¹⁶ An important implication of the high proportion of joint costs relative to user-specific, product-specific marginal costs, is that any attempt to regulate pharmaceutical prices on the basis of costs is at best imprecise and arbitrary, at worst systematically downward biased. The fraction of the joint costs that should be paid by Italians or Americans cannot be determined by accounting rules; rather, the appropriate sharing rule depends on demand conditions in different countries. Thus in the case of pharmaceuticals "costs" do not provide an objective benchmark for setting

The airline industry offers an example of differential pricing that works reasonably well without regulation in an industry characterized by large joint costs and monopolistic competitive market conditions. Since airline deregulation in the US, price differentials have increased while average price levels have fallen significantly. Each airline may have some local monopoly power, but competition between incumbents, reinforced by entry by new airlines, constrains profits to roughly normal levels on average. This may be imperfect, but in such industries the best we can hope for is a rough second best.

6. Compulsory Licensing

Some proposals would permit LDCs to use compulsory licensing of local manufacturers or parallel trade, if such sources can supply products at lower prices than those offered by originators. As argued above, one reason originator firms charge higher prices is the risk of spillovers to other markets, which is not an issue for local firms. The simplest solution is to prevent the cross-national price spillovers, such that originator firms have incentives to price for LDC markets based solely on conditions in these countries.

If under these circumstances, compulsory licensees still have lower prices than originators, and this is not due simply to lower quality, then the lower prices of local firms plausibly reflect lower real costs. If so, there is a case for permitting compulsory licensing of the local generic company and exports to countries that have no local generic producers. This assumes that the benefits to consumers in these LDCs from access to lower price medicines is large, and the revenue loss and hence adverse effect on R&D incentives of originator firms is small because their prices would have approximated marginal cost. However, before adopting this approach, it is important to determine whether the higher prices of originator products are due to either lack of market separability or quality differentials or high distribution margins added by distributors or pharmacies. These problems are best addressed directly, not by compulsory licensing.

The risk of permitting compulsory licensing is that this approach may expand beyond the circumstances where it is justified by real cost advantage for essential drugs for the most needy populations, to a broad range of drugs and countries that seek to use compulsory licensing as a way to avoid making any contribution above marginal cost to pay for R&D. Many middle and even high-income countries face health needs for their populations that exceed the budgets available, as new drugs offer new treatment possibilities for growing elderly populations. It is a fact of life in every country that “needs” are infinite but budgets are finite. Thus many countries could make a hardship case for

prices. If allowable costs are confined to those that can be verified and clearly attributed to a specific product in a specific country, cost-based regulation will lead to prices that are inadequate to cover total costs.

compulsory licensing of a wide range of drugs. In the absence of clear criteria to define which drugs and countries/ populations should be eligible, the compulsory licensing approach is at risk of undermining the function of patents over broad markets and therapeutic categories. This approach may seem to offer cheap drugs to needy people in the short run, but at the risk of undermining incentives to develop new drugs in the longer run. Moreover, even the short run benefits may be illusory since there is no guarantee that generic companies will price at marginal cost. Even if they do, wholesalers and pharmacies often add high distribution margins, exploiting their local monopoly power. Given the risks inherent in the compulsory licensing “solution,” it seems best to first try the approach of strengthening market separation, in order to make it possible for originator firms to maintain differential pricing. In these circumstances originators may offer prices comparable to the prices that a local generic firm would charge, which eliminates the need for compulsory licensing.

A second, often implicit rationale for compulsory licensing is industrial policy, since compulsory licensing has the effect of transferring revenues that might have accrued to a multinational company to a local firm. If there is an implicit infant industry or local production rationale for compulsory licensing, these arguments should be made explicit and evaluated on the merits. This issue is too large to be addressed here but warrants further study.

7. Conclusions

Differential pricing would go a long way towards making drugs that are developed for high income countries available and affordable in LDCs, while preserving incentives for R&D. Differential pricing based on Ramsey pricing principles, which implies prices inversely related to demand elasticities across markets, is consistent with the criterion of economic efficiency. It is also consistent with standard norms of equity.

Unfortunately, actual price differentials are probably not optimal, partly because manufacturers are reluctant to grant low prices in low-income countries because these low prices are likely to spill over to higher-income countries through parallel trade and external referencing.

To achieve appropriate and sustainable price differences will require either that higher-income countries forego these practices of trying to “import” low prices from low-income countries, or that such practices become less feasible. The most promising approach that would prevent both parallel trade and external referencing, is for payers and companies to negotiate contracts that include confidential rebates. With confidential rebates, final transactions prices to purchasers can differ across markets without significant differences in manufacturer prices to distributors, such that opportunities for parallel trade

and external referencing are eliminated. As long as higher income countries can and do attempt to bargain for lower prices that are given to low income countries, companies will rationally be unwilling to grant these low prices to the low-income countries. This severely undermines the ability of these countries to achieve access to existing drugs, which in turn creates hostility to patents. However, patents need not – and probably would not -- entail high price- marginal cost mark-ups in low income countries if companies could be confident that low prices granted to low income countries would not leak to high and middle income countries.

Differential pricing can go a long way towards resolving the problem of making drugs that are developed for high income markets affordable in LDCs, while preserving incentives for innovation. However, differential pricing alone cannot solve the problem of creating incentives for R&D to develop drugs for diseases that are confined to LDCs, since there is no high income market demand to pay for the R&D through higher prices. Differential pricing will also not fully resolve the problems of affordability for existing drugs if these have high marginal costs – due, for example, to high production or distribution costs—or if intermediaries add high margins, such that retail prices are significantly higher than manufacturer prices. Chronic medications, especially those that are costly to produce, may be unaffordable for the neediest populations even at prices close to marginal cost. In such contexts, differential pricing can reduce but not eliminate problem of making drugs affordable to LDC populations.

References

Baumol, W.J. and Bradford, D.F. 1970. “Optimal Departures From Marginal Cost Pricing,” *The American Economic Review*, 265-83.

Breutigam, R.R. 1984. “Socially Optimal Pricing with Rivalry and Economies of Scale.” *Rand J. of Economics*, 127-34.

Danzon, P.M. 1997. “Price Discrimination for Pharmaceuticals: Welfare Effects in the US and the EU,” *International Journal of the Economics of Business*. 4(3): 301-321.

DiMasi, et al., 1991. “The Cost of Innovation in the Pharmaceutical Industry.” *J. of Health Economics* (10):107-42.

Grabowski, H.G., and Vernon, J.M. 1990. “A New Look at the Returns and Risks to Pharmaceutical R&D.” *Management Science* 36(7): 804-21.

Laffont, J-J and J.Tirole, 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge: MIT Press.

PhRMA (Pharmaceutical Research and Manufacturers of America). 2001. *Industry Profile*.

Prieger, J.E. 1996. "Ramsey Pricing and Competition: The Consequences of Myopic Regulation." *J. of Regulatory Economics* (10): 307-21.

Ramsey, F.P. 1927. "A Contribution to the Theory of Taxation," *Economic Journal* (37): 47-61.