Differential Pricing of Pharmaceuticals: Theory, Evidence and Emerging Issues

Patricia M. Danzon

© Springer Nature Switzerland AG 2018

Abstract
Differential pricing—manufacturers varying prices for on-patent pharmaceuticals across markets—can, in theory, lead to increased patient access and improved research and development (R&D) incentives compared with charging a uniform price across markets. Theoretical models of price discrimination and Ramsey pricing support differentials based inversely on price elasticities, which are plausibly related to average per capita income. However, these models do not address absolute price levels and dynamic efficiency. Value-based differential pricing theory incorporates insurance coverage and addresses static and dynamic efficiency. Limited empirical evidence indicates a weak positive relationship between prices and gross domestic product (GDP) per capita. External referencing and parallel trade undermine differential pricing. We discuss previously neglected factors that undermine differential pricing in practice. High price growth relative to GDP in the USA leads to widening differentials between the USA and other countries. Concerns over the effects of confidential rebating challenges acceptance of this approach to implementing price differentials. The growth of branded generics in low- and middle-income countries leads to complex markets with product and price differentiation.

1 Introduction
‘Differential pricing’ is the practice of manufacturers charging different prices for the same product in different markets. In theory, differential pricing between rich and poor countries can increase access to pharmaceuticals in low-income markets while preserving manufacturer revenues and incentives to invest in research and development (R&D) [1–5]. However, the literature shows little consensus on how to implement differential pricing, how to determine appropriate price levels across countries and whether other tools are more likely to increase drug access in low-income countries. The limited empirical evidence finds a generally positive but weak relationship between drug prices and average per capita income across countries. Differential pricing may also occur between payers within a single country—for example, in the USA, different health plans pay different prices for medical services, including pharmaceuticals. The practice, the appropriate price levels and differentials, and implementation through confidential rebates all remain controversial.

1 Retail prices may also differ due to distribution markups and taxes, but these are not discussed here.
2 For example, see Outterson [6] and sources therein.
Differential pricing has strong support in economic theory. When customers differ in their price elasticity, differential pricing can improve access and both static and dynamic efficiency compared with charging the same price to all customers [3]. Static efficiency is improved if charging lower prices to more price-sensitive consumers increases their utilisation of drugs, thereby increasing consumer welfare and profits for producers. Dynamic efficiency is enhanced if differential pricing better aligns prices realised by producers with the value of drugs to consumers, such that producers face appropriate incentives to invest in R&D. Since greater price sensitivity is plausibly associated with lower income, the implication is that prices should differ across countries based on average per capita income. This is also consistent with vertical equity, defined as the wealthy contributing more than the poor to the joint costs of R&D. Although charging different prices for the same product is sometimes viewed as violating horizontal equity or ethical pricing [8], in general there is broad policy acceptance for the principle of cross-national differential pricing, with prices positively correlated with average per capita income or a human development index [2, 7].

Theoretical literature and the policy debate have reached less consensus on specific details of implementation, including absolute price levels, price differentials and enforcement mechanisms. While economic theory provides some support for varying prices according to income, it provides no precise guidance on whether prices should vary less or more than in proportion to gross domestic product (GDP) per capita; whether other factors, such as disease prevalence and budget impact, should be taken into account; whether income or a broader human development index is the best measure of willingness to pay (WTP/ability to pay) [7]; and the role of insurance coverage [3, 9]. In practice, drug prices reflect multiple decisions of manufacturers and payers/customers in each country, and the magnitude of cross-national price differentials is an empirical question.

Section 2 reviews the theoretical foundations for differential pricing. Section 3 summarises recent empirical evidence on prices and known obstacles to implementation, notably external reference pricing (ERP) and parallel trade. Section 4 examines three important—and previously neglected—challenges to differential pricing in practice. First, the divergence of price growth in the USA relative to GDP and price growth in other countries undermines stable price differentials over time. Second, implementing differential pricing through confidential rebates is increasingly common but widely challenged. It is potentially efficient in contexts where either competition or regulation force the pass through of rebates to consumers but can distort pricing and/or prescribing when these conditions fail. Third, the highly skewed distribution of income in many low- and middle-income countries (LMICs) creates incentives for originator monopolists to set high prices targeted to the wealthy minority rather than lower prices based on average per capita income [10]. This phenomenon is exacerbated in markets where originators compete with branded generic substitutes. Such product and price differentiation is more complex than envisaged by standard models of differential pricing, and welfare implications depend on specifics of both pricing and product quality. Section 5 concludes.

2 Theoretical Foundations

Three distinct theoretical approaches have been used to address the optimal pricing of drugs across countries. All support differential pricing, but they differ in assumptions and implications, as described in the following sections.

2.1 Price Discriminating Monopoly

The most general theoretical support for differential pricing is the positive and normative analysis of price discrimination [11, 12]. This theory demonstrates that, if consumers in submarkets differ in their price sensitivity, a monopolist will maximise profits by varying prices across submarkets based inversely on price elasticity, that is, charge lower prices to price-sensitive customer segments than to less price-sensitive segments, assuming submarkets are separable. Under normal conditions, such price discrimination increases utilisation by price-sensitive consumers and, in aggregate, increases consumer welfare and producer profits, relative to charging all customers the same price. True price elasticities (before insurance) are unobservable in practice. However, theory and evidence suggest that price elasticity varies inversely with income because high prices

---

3 In theory, dynamic efficiency requires that producers capture the full expected marginal social surplus created by innovation. This conclusion ignores potential practical issues, for example, that full producer surplus capture may induce excessive ‘racing’ R&D to capture monopoly rents or, if R&D is lumpy, a significant share of surplus may suffice to induce appropriate investment. The underlying theory also assumes that consumers/decision makers on average accurately perceive the benefits and risks of drugs.

4 Differential pricing theory focuses on on-patent products sold by a monopolist with pricing power. Models of markets with generics typically assume that prices are constrained by competition, assuming that the generics are required by regulation to be bioequivalent to the originator, hence quality is known. Such equivalence is not required of branded generics in many LMICs, hence quality uncertainty undermines price competition (see Sect. 4.3).

5 This welfare measure is a simple aggregate of equally weighted consumer utilities.
have a proportionately larger income effect for low-income consumers. This theory supports the conclusion that differential pricing based on average per capita income across countries increases utilisation and enhances both static efficiency and vertical equity, relative to uniform pricing [12, 13]. Under plausible assumptions, differential pricing also leads to higher investment in R&D and improved dynamic efficiency [14].

Thus, economic theory concludes that, even without regulatory constraints, the profit motive leads a monopolist to charge prices across market segments inversely related to price elasticity, and this enhances social welfare compared with charging a uniform price. However, the absolute price levels charged by an unregulated, profit-maximising monopolist may yield above-competitive return on investment (ROI) unless pricing power is constrained by the potential entry of differentiated substitute products and robust consumer price sensitivity, such that the market approximates monopolistic competition rather than pure monopoly [13]. Pharmaceutical markets are subject to entry of differentiated therapeutic substitutes but with entry lags, especially in small markets. More importantly, extensive insurance undermines consumer price sensitivity and can enable prices that yield above-competitive ROI in the absence of regulatory constraints [15].

### 2.2 Ramsey Pricing

A second foundation for differential pricing of pharmaceuticals is ‘Ramsey pricing’ [16]. This was developed as a solution to the normative problem of finding the welfare-maximising (optimal) set of tax rates by which to raise a target revenue, given heterogeneous consumer response to taxes. This is analogous to the problem of finding the optimal set of prices across markets by which to finance pharmaceutical R&D, which entails joint costs that cannot be causally attributed to particular consumers/countries. These costs are sunk at launch, but prices/sales must in aggregate be expected to cover the joint costs to attract continued R&D investment. Ramsey [16] concluded that charging prices that vary inversely with price elasticities is the most efficient way to allocate joint costs, whereas absolute price levels depend on the revenue required.

Thus, both Ramsey pricing and price discrimination support differential pricing across markets based inversely on price elasticities. However, both models ignore insurance, which undermines price elasticities in practice, and both focus on relative, not absolute, prices. Ramsey [16] pricing assumes that absolute price levels are constrained by regulations to raise the revenue required to cover predetermined costs. Ramsey pricing thus addresses optimal pricing for static efficiency, given a predetermined financing requirement. It does not address pricing for dynamic efficiency, that is, to incentivise optimal investment in R&D. By contrast, price-discrimination theory describes the profit-maximising behaviour of an unregulated monopolist and shows that the unregulated monopolist is induced by profit incentives to charge price differentials similar to the Ramsey-optimal differentials. However, unconstrained monopoly prices may yield an above-competitive ROI and thus fail to achieve either static or dynamic efficiency unless constrained by monopolistic competition and robust consumer price sensitivity, which in pharmaceutical markets is undermined by comprehensive insurance [13, 15].

### 2.3 Value-Based Differential Pricing

Value-based differential pricing (VBDP) is a third approach to implementing differential pricing that is designed to achieve both static and dynamic efficiency in markets with comprehensive insurance, as exists in most high-income countries (HICs). In such contexts, payers set rules for prices/reimbursement and utilisation that can be designed to achieve both static and dynamic efficiency, that is, the broadest possible utilisation of existing drugs and R&D incentives to develop new products that deliver expected benefits that justify expected costs [3]. The basic principle is that prices in each country should reflect its WTP for health. In countries with comprehensive insurance, this can be achieved if each public or private payer, acting as agent for its enrollees, sets a WTP threshold, defined as a maximum cost per unit of health [e.g. cost per quality-adjusted life-year (QALY) or other metric] that new technologies must meet as a condition of reimbursement. Specifically, each payer should evaluate whether the incremental cost effectiveness of a new drug, relative to comparator treatments, meets that payer’s WTP threshold, given the evidence on the drug’s incremental benefit, its price and any related costs and offsets. Making reimbursement of new technologies contingent on meeting a WTP threshold creates incentives for manufacturers to set the price for a drug commensurate with its incremental value (including health gain and cost offsets) and to invest in R&D for products that can meet the value thresholds. Thus, a new drug that is highly effective relative to its comparator can

---

6 Overall price elasticity combines an unobserved pure (income-compensated) price elasticity and an income effect due to the price change. This income effect is expected to be positive. For detail, see Danzon et al. [3].

7 R&D investment to establish safety, efficacy and manufacturing standards for new drugs entails joint costs and creates a knowledge base that can benefit consumers globally.

8 WTP in poor countries could include payment by citizens and donors.
be priced at a significant premium, whereas a new drug that offers no incremental benefit must price at par in order to be reimbursed.

From a global perspective, if each country/health plan unilaterally sets its own threshold WTP for health, this creates incentives for manufacturers to set price differentials across countries/plans that reflect their respective WTP for health. The aggregate sales based on such prices reflects global WTP and creates optimal incentives for R&D.

Similar to price discrimination and Ramsey pricing, VBDP implies WTP thresholds and price differentials across countries that are plausibly positively related to income and other factors. In contrast to price discrimination and Ramsey pricing, VBDP in theory provides an approach to setting actual prices to achieve both dynamic and static efficiency.9 VBDP recognises the role and effects of insurance and relies on payers to set reimbursement rules based on a WTP approach that is already approximated in some countries. It does not require a supra-national regulator for implementation but assumes unilateral reimbursement decisions by payers.

### 2.4 Imperfect Market Separation: External Referencing and Parallel Trade

In practice, pharmaceutical pricing is implemented through some combination of markets and national price/reimbursement regulation in most countries. Maintaining price differentials across countries requires that markets are separate. In fact, two specific policies adopted in many countries enable regulators and/or middlemen to arbitrage ex-manufacturer price differences to varying degrees, which in turn influences manufacturer strategies [12–14]. First, because free trade is the norm for most goods, especially within regional trading blocs, market separation requires that countries ban ‘parallel trade’ of on-patent drugs and adopt national exhaustion of patents.10 By contrast, the EU upholds parallel trade of drugs within the EU (but not from outside the EU) and has adopted a rule of ‘community exhaustion’ of patents. Thus, once a patent holder sells a good anywhere within the EU, it can be resold by distributors elsewhere within the EU. Distributors’ rights to conduct parallel trade have repeatedly been upheld by the European Court of Justice.11

Second, many payers use ERP relative to prices paid in other countries as one benchmark in setting their domestic prices. Such ERP intentionally constrains differential pricing, capping domestic prices at the mean, median or minimum price in the basket of referenced countries. ERP has been adopted as a price/reimbursement tool by all EU countries except Sweden and the UK [17]. ERP appeals to policy makers because it requires relatively simple and objective data, and the benchmark countries and price point (median, minimum, etc.) can be selected to put downward pressure on domestic prices. By contrast, value-based pricing and related cost-effectiveness analysis require data on incremental outcomes and costs, relative to comparators, which can entail costs and contention.

Several studies confirm theoretical predictions that parallel trade and ERP contribute to companies’ reluctance to sell at low prices, especially in small, lower-income countries in the EU, because low prices in any EU country can undermine potentially higher prices in other EU countries [18–20]. ERP has been adopted in some LMICs and is under consideration in others. Although most countries chose reference countries that are regionally close and/or at similar levels of income, the widespread use of ERP and parallel trade plausibly contributes to originator companies’ reluctance to set low prices even in low-income countries. Vaccines are an exception with sustained differential pricing, plausibly because the vaccine distribution chain from manufacturer to final provider is more tightly controlled to assure product safety and prevent product diversion to unintended markets.

### 3 Empirical Evidence on Cross-National Differential Pricing

This section reviews empirical studies that focus on measurement of cross-national price differences for drugs. We sought studies that compared prices for comparable drugs across many countries, with consistent methodologies over time, to track how price differentials were related to per capita income and how this relationship changed over time. In practice, the small number of academic studies of cross-national price differentials were generally not comparable because of differences in countries, drugs and methodologic issues used, including non-comparable formulations (tablets/capsules/liquids) and pack sizes; price measurements (per daily dose, per gram or standard unit); and weighting used to calculate average price differences; and whether exchange rates or purchasing power parities were used for currency conversion. Further, although differential pricing theory applies only to on-patent originator drugs, some empirical studies include generics.

---

9 Consistent with the patent system and other policies, the aim here is second-best efficiency, recognizing that implementing first-best efficiency would entail taxes and subsidies that themselves entail administrative costs and are therefore not generally considered practical.

10 Parallel trade is also called ‘commercial drug importation’ in the USA.

11 Manufacturers may limit supply to each country to the quantity needed by that country, without such restriction constituting a boycott.
Differential Pricing of Pharmaceuticals: Theory, Evidence and Emerging Issues

Fig. 1 Average foreign-to-Canadian price ratios, patented drugs: Organisation for Economic Co-operation and Development (OECD), 2016 [36, 51]. The y axis shows the average (arithmetic mean) of foreign-to-Canadian price ratios for individual drugs, based on retail and hospital sales, with Canadian sales weighting, using IMS data. Off-invoice rebates are not reflected, but US prices incorporate prices from the Federal Supply Schedule (FSS). IMS currency data converted at market exchange rates [36]. GDP gross domestic product

Fig. 2 Average foreign-to-Canadian price ratios, 2005, 2016. The Patented Medicine Prices Review Board requires patentees to report publicly available ex-factory prices in seven comparator countries. The US price incorporates prices from the Federal Supply Schedule (FSS), which reflects some mandatory and negotiated rebates. Price indexes are Canadian sales-weighted arithmetic means of price ratios for individual drug products. Currency conversion at market exchange rates [36].
Studies of drug prices across predominantly HICs generally find a positive correlation between prices and income, but reimbursement rules and other factors clearly also influence both drug pricing and availability (Fig. 1) [22–24, 27, 36]. The excess of US prices over those in other countries has increased over time (Fig. 2) [27, 36]. Within the EU, parallel trade and ERP constrain differential pricing and reduce access in lower-income countries, as companies choose delay or non-launch rather than accept low prices that could undermine higher-price markets [18–20, 26]. Parallel trade has not necessarily reduced prices in importing countries, because middlemen incur costs and capture some of the spread [25]. Obstacles to differential pricing in the EU have increased with the accession of Eastern European countries and the growing use of ERP, often with an EU-wide reference basket [17, 26]. The spillover relationships are sufficiently complex that manufacturers may rationally choose to price within a narrow corridor of relatively high prices, preferring delay or non-launch in countries that cannot meet price targets.

Accuracy measurement of transaction (net) prices has become increasingly problematic due to the growing use of confidential rebates and other risk- or cost-sharing measures agreed between payers and manufacturers [26]. These include product-specific rebates (the payer receives a rebate if patients do not achieve target outcomes) and broader provisions for price rollbacks or rebates if drug expenditures exceed targets. Such confidential/off-invoice rebates reduce costs for payers and are preferred by manufacturers to list price reductions that would spill over to other countries through ERP. Thus, confidential rebates can facilitate differential pricing and promote access in countries/plans using such approaches. Although other countries/payers complain that confidentiality undermines transparency and their ability to practice ERP, this is not problematic per se, given that ERP is welfare inferior to differential pricing, assuming that all confidential rebates are passed on through to final payers (see Sect. 4).

For LMICs, the limited empirical evidence shows relatively high originator prices relative to GDP per capita and a weak correlation between prices and average income across countries [21, 28–30]. High originator prices in low-income countries cannot be explained solely by fears of ERP or parallel trade, because such practices are limited and occur mainly between LMICs at similar income levels. Other factors include the highly skewed income distribution in many LMICs, which makes it more profitable for a monopolist to set high prices targeted to the high-income segment than to set lower prices that could be affordable to the less-affluent majority [10]. Competition from branded generics reinforces this segmentation incentive for the originator and reduces its negative welfare effects on consumers [21]. We return to this in Sect. 4.3.

4 Emerging Issues in Differential Pricing

Previous sections reviewed the underlying theory and empirical evidence on differential pricing across countries, including ERP and parallel trade, that undermine necessary market separation and are commonly blamed for the failure of differential pricing. This section discusses three issues that have been largely ignored by previous analyses but are increasingly important in understanding why differential pricing fails in practice.

4.1 Divergent Rates of Price Growth: US vs. Ex-US

In recent years, pharmaceutical price growth has diverged between the USA and other HICs due to differences in their insurance and payer strategies. Most HICs have comprehensive universal health insurance programs that seek to maximise health for citizens within annual health budgets set at a stable percentage of GDP. Payers review prices of new technologies to ensure any price premium over established treatments is justified by incremental benefits. Post-launch price increases are disallowed unless supported by evidence of additional benefits [31, 36].

By contrast, in the USA, most public and private payers face looser budget constraints, raising premiums or funding requirements (Medicare) to accommodate spending increases.12 Manufacturers set drug prices freely, at launch and with post-launch increases [32]. This pricing freedom has enabled originator drug price growth that exceeds general inflation due to both rising launch prices and post-launch price increases. Median launch prices for new cancer drugs increased from $US100 to $US10,000 per month of treatment from 1960 to 2016 [33]. Adjusting for benefit gain, the cost per life-year saved rose on average by $US8500 per year since 1995 [34]. For orally administered cancer drugs, the average cost per month increased from $US1869 in 2000 to $11,325 in 2014, after adjusting for inflation [35]. This excess of drug price growth over general inflation is not confined to oncologics. The aggregate Pharmaceutical Producer Price Index (PPI) increased 83% from 126.8 in January 2007 to 231.5 in December 2016, with the annual growth rate increasing from 4.1% in 2007 to 8.8% in 2016.13 This aggregate index understates price growth for originator drugs because it includes generics.14

12 Medicaid is an exception, where states must operate within annual budgets.
14 The Bureau of Labor Statistics PPI treats generics as new forms of originator compounds, hence the price of the compound drops significantly following the entry of cheap generics. Generic prices generally decline post-launch [37, 38].
Between 2008 and 2016, the Express Scripts Price Index for Brand-Name (originator) Prescription Drugs increased threefold, whereas the Generic Prescription Price Index fell over 50% [38]. For originator drugs in the USA, annual price increases of up to 10% are now a norm. Annual reports from Canada’s Patented Medicine Prices Review Board (PMPRB) show the contrast between positive annual price growth in the USA and flat/negative price change in major EU markets and Canada [36].

High US drug price growth reflects the traditional reliance of US payers on patient cost sharing and price sensitivity to constrain pricing. However, patient cost sharing has become ineffective, as insurance now covers over 85% of total drug spending, drugs are increasingly differentiated and prices exceed the annual ‘stop-loss’ limit on patient cost sharing. Stop-loss limits provide essential financial protection for consumers but inevitably undermine patients’ price sensitivity beyond the limit. Thus, for expensive drugs, patients who lack insurance cannot afford them, and insured patients are price insensitive. Medicare is barred by law from negotiating prices, and private payers can, at most, negotiate rebates off list prices in drug classes with close therapeutic substitutes.

This divergence of price trends between the USA and other HICs implies that price differentials roughly proportional to per capita income, which roughly prevailed in the late 1990s, are unsustainable, and increased divergence is likely. The Canadian PMPRB reports that the US/Canada price index for on-patent drugs increased from 1.83 to 3.08 between 2005 and 2016, based on manufacturer data supplied to the PMPRB. By contrast, price indexes for six major EU countries versus Canada remained stable, ranging from 0.77 to 1.21 (see Fig. 2), implying that US prices also diverged relative to these major EU markets. Figure 1 shows 2016 price indexes for all Organisation for Economic Co-operation and Development (OECD) countries relative to Canada, using IMS MIDAS data [36]. Relative to Canada, US prices are highest, at 2.91, followed by Switzerland at 1.06, Germany at 1.00 and the OECD median at 0.80. Because these data omit off-invoice rebates, US prices may be overstated relative to foreign prices, assuming rebates are more pervasive in the USA. However, no reasonable allowance for omitted rebates could bring these price differentials in line with GDP differentials, which are also shown in Fig. 1. The relationship between prices and income is positive but weak. The US price index is 3.64 times the OECD median price index, whereas US GDP is only 1.38 times median OECD GDP per capita. Thus, US drug price differentials cannot be ‘explained’ by income differentials.

This divergent growth of US drug prices implies increasing strain on reimbursement negotiations for new drugs ex-USA if manufacturers seek to maintain income-related differentials relative to their US prices, because price differentials may be cited in support of policy proposals in the USA for ERP and/or drug importation. However, for payers ex-US, income-based differentials applied to rising US launch prices would be out of line with their domestic prices for existing comparator drugs. This leads to increasing risk of rejection and/or confidential rebates.

To illustrate, assume that drug X was launched at the same price in the USA and Germany in 2008, that annual originator price growth in the USA and Germany has been 10 and 0%, respectively, and exchange rates have not changed. By 2018, the US–Germany price differential for X would exceed 150%. If drug X is now used as comparator in both countries for pricing new drug Y, which has a small incremental benefit, the excess post-launch price growth in the USA implies a spread of over 150% between the prices that the manufacturer and German payer, respectively, would deem a value-based price, based on Y’s incremental benefit relative to X. If a country has maintained a roughly stable cost-per-QALY threshold for reimbursement, as in the UK, manufacturer pricing of new drugs based on US benchmarks will increasingly exceed this cost-effectiveness threshold unless manufacturers are willing to accept widening price differentials relative to the USA. In practice, this tension driven by divergent price growth is resulting in widening US–EU list price differentials [36] and increased use of confidential rebates off list price through ‘patient access schemes’ that reduce net price to payers while limiting the spillover risk to manufacturers through ERP or parallel trade [26].

4.2 Confidential Rebates: Efficient or Distorting?

In the USA, pharmacy benefit managers (PBMs) manage outpatient drug coverage for most private plans, and similar prescription drug plans (PDPs) manage Medicare Part
D. A key tool of PBM cost management is to negotiate with manufacturers for rebates off list prices in return for preferred formulary placement designed to increase market share [13, 42]. The rebates are confidential and paid through electronic transfer from the manufacturer to the PBM, thereby preempting wholesalers or pharmacies from capturing the price spread or diverting the discounted product to unintended customers. Rebates may be contingent on a drug’s actual market share to incentivise PBMs to employ other management tools to shift utilisation towards preferred drugs. Confidential rebating enables manufacturers to sell drugs to wholesalers at a common list price while offering rebates directly to those PBMs/plans that adopt strategies to increase a drug’s market share. Such rebating can be an efficient form of price competition provided that competition between PBMs forces them to pass through rebates as lower premiums to plans/encelees who accept restrictions on their drug choice in return for lower premiums.

However, critics charge that rebates are not fully passed through and that confidentiality enables PBMs to retain rebates, which distorts their incentives to design the most cost-efficient formularies for patients. In theory, rebate confidentiality encourages price competition between manufacturers, whereas rebate transparency would encourage tacit collusion on price. The empirical evidence on the pass through of PBM rebates is limited and inconclusive, partly because PBM cost structures, revenue sources and contracts with employer sponsors are complex [42]. The increased concentration of the PBM industry, combined with the intrinsic difficulties faced by sponsors in evaluating PBM performance, creates concern over whether the competitive pressures faced by PBMs are sufficient to force full rebate pass through to plan sponsors and lower premiums for enrollees.

Recent proposals that patient cost sharing should be based on the net-of-rebate price rather than list price [50] raise additional issues. Basing cost sharing on net-of-rebate prices would require that drug-specific rebates be transparent to not only sponsors/patients but also the dispensing pharmacies. PBMs that own and operate pharmacy chains would thus gain an anti-competitive advantage because they would observe the rebates obtained by competitor PBMs. This could encourage further consolidation between PBMs and pharmacy chains and discourage rebating and would not address the underlying problems.

If weak competitive pressure enables PBMs to retain rather than pass through some share of drug rebates, then PBMs have incentives to prefer high-price/large-rebate drugs over competitor drugs with lower prices/lower rebates/lower net price to sponsors, as is sometimes alleged. If so, confidential rebating encourages manufacturers to raise prices and rebates, and the system distorts rather than promotes efficient competition and lower prices. Further, rebate contracts may be structured to deter entry and be subject to antitrust challenge in both health and non-health contexts [43].

Some in the pharmaceutical industry blame PBMs for high drug prices, arguing that rebate demands by PBMs force manufacturers to raise list prices [45]. This argument implicitly acknowledges that list prices can be increased without effective limit. This is unsurprising, since Medicare is barred by statute from negotiating prices, and private plans lack market power over list prices. Given this pricing freedom, economic theory and recent evidence of price increases [33–36, 38, 39, 46] suggest that manufacturers would rationally raise prices to increase their revenues, beyond offsetting rebates. Thus, confidential rebating off list price fails as an efficient, differential pricing mechanism in circumstances where manufacturers can raise list prices without constraints and intermediaries face weak competitive pressures to pass through rebates to ultimate payers. However, in such contexts, rebate transparency would not halt the rise in list prices and could discourage competitive rebating.

The use of confidential rebates to implement differential pricing within and between countries is growing ex-USA, together with concern about whether such practices improve or harm social welfare. In general, discounts/rebates can be consistent with efficient, competitive differential pricing,

---

20 PBMs are intermediaries/agents contracted by health plans, self-insured employers and other plan sponsors to manage drug coverage. Some very large health plans manage their own PBMs [42].

21 PBMs’ utilization management tools include very restricted formularies, with only one or two preferred drugs per class; a large spread in patient copayments for preferences versus non-preferred drugs; step edits; and prior authorization for non-preferred drugs.

22 The theory that price transparency facilitates collusion by competitor firms was developed by Stigler [40]. Consistent with this, the Congressional Budget Office estimated higher costs for Medicare Part D if PDPs are required to publish drug-specific rebates [41].

23 For example, two dominant, incumbent drugs in a class could create barriers to entry of competitors if they make their rebates contingent on being one of only two preferred drugs in the class for a formulary. If a PBM were to add a new drug on the preferred tier, either in addition to or in place of an incumbent, because the new drug offered a lower list and net price, the PBM could lose significant rebate revenue from incumbents, especially if the uptake of the new drug is slow due to prescriber/consumer brand loyalty to the incumbent drugs.

24 For example, Graf [44] examined the rebate contracts used by German Sickness Funds, which are now authorized to contract with manufacturers on behalf of their enrollees using exclusive or non-exclusive contracts. In several Asian countries, drug dispensing by providers has encouraged rebating by manufacturers to providers. Recent policies in South Korea, Japan and China seek to discourage provider dispensing and manufacturer rebating as potentially harmful to efficient drug prescribing.
provided that list prices are constrained and rebates passed through to ultimate payers/customers. Discounts/rebates can distort efficiency when they benefit intermediaries or providers, particularly when list prices are also unconstrained. Rebates can also be primarily distributional, with uncertain effects on efficiency, if they simply reflect relative bargaining power in a bilateral monopoly.  

4.3 Low- and Middle-Income Countries: Differential Products and Prices

Many LMICs have high originator drug prices relative to GDP per capita, contrary to the theoretical prediction that monopolists would set prices roughly commensurate with income as a proxy for price elasticity. High prices limit consumer access in LMICs, where most patients lack insurance coverage for outpatient drugs.

Flynn et al. [10] showed that, faced with the highly skewed income distribution typical of LMICs, a monopolist would maximise profits by charging a relatively high price targeted to the wealthy rather than a more affordable price for average-income consumers. In theory, a monopolist could increase profits by also offering a differentiated formulation priced for lower-income consumers, either directly or by licensing a generic producer. However, originators rarely adopt such strategies, partly because selling a cheaper formulation could cannibalise sales of the higher-priced formulation if the products are known to be the same.

In practice, middle- and lower-income consumers in LMICs are served by branded generics, which claim similar therapeutic effects but in many countries are not required to meet regulatory standards of bioequivalence [37, 48]. Since quality of branded generics is uncertain, they compete on brand, and price becomes a proxy for quality. Domestic-branded generic firms are numerous in countries such as India, China, Brazil, Mexico and South Africa. Some large generics producers have achieved regional and/or multinational status as suppliers that can meet World Health Organization (WHO) quality requirements, and US FDA requirements when selling to the USA, for both biologic and chemical drugs [21, 49]. Originator multinational corporations (MNCs) may rationally choose not to cut price to compete with these branded generics, because (1) cutting price to lower-income consumers could ‘cannibalise’ sales at higher prices and risk spillovers to other countries; (2) MNCs have higher cost structures than branded generics; and (3) serving large segments of the huge LMIC populations would likely require additional investment in manufacturing capacity with uncertain ROI.  

Thus, the norm in LMICs is product and price differentiation, with originators charging high prices to price-inelastic, quality-inelastic high-income customers, whereas branded generics offer lower-priced, less-certain quality alternatives to lower-income/price-sensitive customers. However, this outcome is likely welfare superior to the alternative, with only a single originator product priced to target the wealthy, with no affordable alternatives for the less affluent, provided that the branded generics are of adequate quality. It is potentially also welfare superior to the alternative in which the originator licenses a single local generic firm to produce a bioequivalent product at a lower price, which would assure product quality but with more limited price competition and affordability, depending on the terms of the license.

Thus, compared with the alternative of a single originator product targeted to the wealthy, adding branded generics to serve the excluded consumers offers potentially significant static efficiency gain and little if any dynamic efficiency loss. Although the traditional differential pricing model, with originator products priced to target average income consumers in LMICs, appears to fail in practice, this model ignores the complexities and opportunities in these markets. Given the distribution of population and income in these countries, the emergence of branded generics to serve customer segments that are unserved by originator products may improve consumer access and static efficiency, with little if any harm to dynamic efficiency. However, conclusions may differ across countries and product types, and further empirical analysis is needed.

5 Concluding Comments

Differential pricing across countries, based roughly on GDP per capita, remains a potentially efficient strategy that could increase consumer and producer welfare, relative to charging a uniform price across all countries. This model is most easily implemented in countries with a universal national or social insurance system where a single purchaser/regulator reflects average citizen preferences and WTP for health. In countries with pluralistic payers, confidential rebating off a single list price can, in theory, achieve competitive, differential pricing between

---

25 A study of insurer/provider contracting in the USA found that large insurers and small provider groups obtained relatively low prices [47], as predicted if differential pricing primarily reflects relative bargaining power. Further research is needed to evaluate the efficiency implications, if any, in such contexts.

26 In theory, MNCs could outsource manufacturing to contract manufacturers, but the MNC would still incur the cost of capacity construction. Agency issues may also be best handled if the LMIC manufacturer also markets the product under its own name and handles distribution, regulatory and marketing, as is currently the norm, rather than contracting with the originator firm to produce its product under license.
payers. Such confidential rebating is potentially efficient provided that list prices are constrained and rebates passed through to ultimate payers/consumers. If list prices are not constrained and/or rebates are captured by agents rather than passed through to final payers, such rebating may contribute to higher list prices and distorted prescribing incentives.

By contrast, in LMICs where most consumers pay out of pocket for drugs, simple differential pricing appears to fail in that the optimal pricing strategy for originator products is to target the high-income segment, not the average-income consumer. However, lower-income segments are increasingly served by lower-priced, branded generics, yielding a differential products and pricing outcome that increases access, aggregate utilisation and static efficiency, with little effect on dynamic efficiency.

Thus, the basic insights of differential pricing remain relevant but with important modifications based on recent experience. First, stable, cross-national income-related differentials are undermined by divergent price growth in the USA. Second, use of confidential rebates is efficient only if list prices are constrained and rebates passed through to ultimate payers/consumers. Third, product and price differentiation in LMICs plausibly achieves broader consumer access and is likely welfare superior to differential pricing with a single originator product. However, this conclusion is tentative, depending on the quality of the differentiated products.

Compliance with Ethical Standards

Funding No sources of funding were used to conduct this study or prepare this manuscript.

Conflicts of Interest Patricia M. Danzon and the Celia Moh Foundation have no conflicts of interest that are directly relevant to the content of this article.

References


