

NBER WORKING PAPER SERIES

VALUE-BASED DIFFERENTIAL PRICING:
EFFICIENT PRICES FOR DRUGS IN A GLOBAL CONTEXT

Patricia M. Danzon
Adrian K. Towse
Jorge Mestre-Ferrandiz

Working Paper 18593
<http://www.nber.org/papers/w18593>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2012

This research was supported in part by Eli Lilly Inc.'s Project on Fair Prices for Pharmaceuticals. We would like to thank the sponsor for this support. The conclusions and views expressed here are not necessarily those of Eli Lilly Inc, nor of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w18593.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Patricia M. Danzon, Adrian K. Towse, and Jorge Mestre-Ferrandiz. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Value-Based Differential Pricing: Efficient Prices for Drugs in a Global Context
Patricia M. Danzon, Adrian K. Towse, and Jorge Mestre-Ferrandiz
NBER Working Paper No. 18593
December 2012
JEL No. I13,I18

ABSTRACT

This paper analyzes pharmaceutical pricing between and within countries to achieve second best static and dynamic efficiency. We distinguish countries with and without universal insurance, because insurance undermines patients' price sensitivity, potentially leading to prices above second-best efficient levels. In countries with universal insurance, if each payer unilaterally sets an incremental cost effectiveness ratio (ICER) threshold based on its citizens' willingness to pay for health; manufacturers price to that ICER threshold; and payers limit reimbursement to patients for whom a drug is cost-effective at that price and ICER, then the resulting price levels and use within each country and price differentials across countries are roughly consistent with second best static and dynamic efficiency. These value-based prices are expected to differ cross-nationally with per capita income and be broadly consistent with Ramsey Optimal Prices. Countries without comprehensive insurance avoid its distorting effects on prices but also lack financial protection and affordability for the poor. Conditions for efficient pricing in these self-pay countries include that consumers are well-informed about product quality and firms can price discriminate between rich and poor subgroups within and between countries.

Patricia M. Danzon
Health Care Management Department
The Wharton School
University of Pennsylvania
3641 Locust Walk
Philadelphia, PA 19104
and NBER
danzon@wharton.upenn.edu

Jorge Mestre-Ferrandiz
Office of Health Economics
Southside, 7th Floor
105 Victoria St.
London SW1E 6QT
United Kingdom
jmestre-ferrandiz@ohe.org

Adrian K. Towse
Office of Health Economics
Southside, 7th Floor
105 Victoria St.
London SW1E 6QT
United Kingdom
atowse@ohe.org

I. Introduction

Achieving efficient pricing of pharmaceuticals between and within countries is a complex conceptual and policy problem. In any industry, pricing to maximize social welfare must consider both static efficiency (optimal use of existing products) and dynamic efficiency (optimal investment in R&D). Reconciling these objectives is problematic for pharmaceuticals, for three reasons.

First, R&D is roughly 17 percent of sales for the US-based pharmaceutical industry, compared to 4 percent for other US industries, and other quasi-fixed costs of production are significant. Marginal cost pricing to achieve first best static efficiency would fail to cover total costs and would violate the dynamic efficiency requirement, that producers capture the full social surplus produced by innovation. Patents enable firms to price above marginal cost and thus potentially achieve dynamic efficiency. This would be “second best” to the extent that pricing above marginal cost reduces utilization.

Second, in the case of pharmaceuticals the effect of patents is both mitigated and distorted by insurance coverage. Most industrialized countries have universal coverage with modest patient co-payments. By lowering out-of-pocket prices to patients, insurance potentially brings utilization closer to first best optimal levels. It also protects consumers from financial risk and, through cross-subsidies, makes health services more affordable to low-income consumers. However, because such insurance makes patient demand highly price-inelastic, insurance creates the potential and incentives for manufacturer prices that exceed the level that would result from patents alone. Public and private insurers use various price control strategies to constrain this producer moral hazard but these controls are generally ad hoc and not designed to achieve static and dynamic efficiency. By contrast, many middle and lower income countries (MLICs) lack comprehensive insurance. Patients in such self-pay markets lack the financial protection of insurance but also avoid its distorting effects on prices. However, other factors – including uncertain product quality and skewed income distributions -- contribute to higher drug prices than might occur in well-functioning markets [1, 2].

Third, because R&D is a global joint cost benefiting consumers worldwide, efficient global pricing requires appropriate contributions from different countries to this joint cost.

Economic theory suggests that price discrimination across countries is welfare superior to uniform pricing if price discrimination increases drug utilization, which is plausible given income differences. Ramsey pricing theory also supports differential pricing across countries, as the second best efficient way to recoup R&D cost, assuming that demand elasticities differ across countries. However, both price discrimination and Ramsey pricing theories give criteria for an efficient structure of *relative* prices but do not address optimal *absolute* price levels or suggest a practical strategy for implementation.

This paper develops a framework of analysis and assumptions designed to achieve efficient absolute and relative prices between and within countries for pharmaceuticals. Although the focus here is on pharmaceuticals, the framework could be applied to other health technologies. The challenges and appropriate solutions in pharmaceutical pricing differ, depending on whether a country has comprehensive insurance. We therefore contrast two different insurance contexts: first, a market with universal coverage (relevant to most industrialized countries) and second, a predominantly self-pay market, where consumers pay cash out-of-pocket for drugs (relevant to many MLICs). We call the set of efficient prices that result from our framework “value-based differential pricing.” The UK Office of Fair Trading use of the term “value-based pricing” to describe a “system that would relate the prices of products to their clinical value relative to existing treatments”[21]. We extend this concept by designing an approach to not only relate prices to value but also maximize health gain from a given budget in each country and achieve static and dynamic efficiency globally.

In countries with comprehensive insurance, pharmaceutical pricing is complicated by potential supplier and consumer moral hazard. In this context, we show that, given certain assumptions, second best optimal prices and utilization can be achieved if each payer or country defines an incremental cost effectiveness ratio (ICER) threshold that reflects its citizens’ willingness to pay for health gain, and pays for drugs that meet this value-for-money standard. These ICERs may differ by payer and/or drug or patient type (e.g. drugs for rare diseases or terminal medical conditions). We assume that manufacturers can set their prices freely, albeit constrained by their drug’s incremental effectiveness relative to the comparator current treatment. We assume that payers limit reimbursement to patients for whom the drug is cost-effective at the manufacturer’s chosen price, in which case utilization will be second best optimal. If each country unilaterally sets its ICER threshold(s) to reflect its willingness-to-pay

for health gain, manufacturer price levels will differ appropriately across countries, and aggregate global revenues will create appropriate incentives for R&D. The ICER thresholds and resulting price levels can be expected to vary across countries in positive relation to per capita income, but the optimal relationship – prices varying more or less than in proportion to income – cannot be determined *a priori*.

Self-pay countries without comprehensive insurance (many MLICs) avoid the distorting effects of insurance on manufacturer pricing, but also forego the insurance benefits of assuring financial protection and affordability of appropriate drugs for patients. Efficient pricing in self-pay markets requires that: regulatory systems assure that all drugs, including for generic copies, meet specified quality standards; consumers or their agents are reasonably informed about drug quality and prices; and manufacturers can price-discriminate within countries that have highly skewed income distributions. In practice, these conditions are not met in many MLICs, leading to pricing above marginal cost and static efficiency loss that potentially outweighs the contribution of high prices to profits and dynamic efficiency [1, 2].

Previous literature has addressed components of a theory of efficient drug pricing within a single country [3–7] and efficient price differentials across countries [8-12]. We contribute to this literature by outlining a comprehensive approach to simultaneously address within-country absolute price levels and between-country relative prices, distinguishing insured vs. self-pay markets. Our analysis is relevant for drugs that treat global diseases. It does not address “neglected disease” drugs, for which global demand is insufficient to incentivize private R&D. In the remainder of this paper, section II reviews previous literature; section III develops the value-based differential pricing (VBDP) model for countries with comprehensive insurance; section IV considers self-pay countries; section V discusses how VBDP differs from Ramsey pricing; section VI concludes.

II. Previous Literature

Optimal Use of Medical Care, No Insurance, Exogenous Prices and Technologies

Garber and Phelps (GP) [3] develop a model of an uninsured individual’s optimal allocation of a fixed budget between medical care and other services, treating the availability and price of medical technologies as exogenous. At the optimum, the patient uses each medical service such that its ICER is equal to the patient’s marginal benefit in terms of expected future

utility normalized by the current marginal utility of income, or willingness-to-pay. GP briefly discuss the application of this model to third party payers, but do not address how a payer might use it. We extend the model to the context of a payer seeking to optimally constrain prices and utilization when faced with price-insensitive patients but endogenous prices and R&D.

Using Insurance to Achieve Static and Dynamic Efficiency in a Single Country

A series of papers [4-5, 7] propose using a standard insurance two-part pricing structure of consumer co-payment and payer top-up payment to reconcile static and dynamic efficiency in a single country. To achieve static efficiency, they set co-payments equal to marginal cost and assume that consumers choose first best utilization. Payers pay top-up payments to transfer the social surplus to manufacturers, as required for dynamic efficiency.

Garber et al. [4] examine the optimal coinsurance rate and incentives for innovation when a monopoly firm pursues profit maximizing pricing. Lowering the coinsurance rate encourages utilization, but also makes demand more inelastic and hence leads firms to charge higher prices. They show that with an unconstrained monopoly price, the optimal coinsurance rate for static efficiency may result in monopoly profits exceeding consumer surplus. They conclude that limits on monopoly pricing may be necessary to avoid excessive incentives for innovation, but do not discuss how to set such price limits.

Lakdawalla and Sood (LS) [7] argue that subsidized but optional public drug insurance with unregulated manufacturer prices, consumer choice of drugs, and a parallel uninsured market, can in theory achieve both static and dynamic efficiency. In their model, static efficiency is achieved in the public program by setting the consumer co-payment equal to the manufacturer's marginal cost. To achieve dynamic efficiency, the insurer tops up this co-payment to the price the manufacturer charges to uninsured consumers. To deter a manufacturer from pricing above the uninsured monopoly p_m , the public payer structures co-payments such that a \$1 price increase above p_m increases the co-payment by more than a \$1. They note that in practice the public payer might constrain manufacturer price below p_m to achieve short run savings, but that this would undermine dynamic efficiency.

This model offers a theoretically elegant solution, but several key components are problematic. First, defining marginal cost in an industry with multiple fixed and quasi-fixed costs poses conceptual ambiguities, specifically, whether the co-payment should reflect just the

(usually minimal) incremental cost of one more dose, or also include an allocated share of quasi-fixed/joint costs of plant capacity and country-specific overhead other than R&D. Further, models of optimal drug co-payments should also consider effects on use of substitute and complementary medical services, notably physician and hospital services [14]. Second, with universal public or social insurance, as exists in most industrialized countries, there is no uninsured monopoly price p_m which the payer can use to benchmark top up payments. Even in countries like the US with an observable price to uninsured patients, this does not necessarily reflect the theoretically ideal monopoly price p_m . In the US, for example, the price manufacturers charge to uninsured patients plausibly exceeds p_m because this uninsured price is used by private and public insurers as the starting point for setting discounts and rebates, which creates incentives to raise p_m . Third, LS assume that raising funds to subsidize public insurance is costless. If taxation in fact entails significant deadweight loss (see, for example, [13]), such that paying top-up payments to manufacturers is not a costless transfer, aiming for second best utilization and top-up payments may be preferable to first best.

Jena and Phillipson (JP) [5] discuss the theoretical relationship between average endogenous cost-effectiveness, consumer and producer surplus and incentives for innovation. They assume that manufacturer pricing is driven by a consumer demand schedule that reflects patients' true marginal benefit, regardless of their insurance coverage. Thus if payers use an ICER threshold as a condition of reimbursement, manufacturers adapt their prices to meet the threshold. JP show that a reduction in price that increases utilization may increase static efficiency but decrease the ex post, average endogenous cost-effectiveness ratio, if the additional users have lower incremental benefit. They do not consider how payers might use an exogenously determined ICER threshold to achieve dynamic efficiency.

Optimal price differentials across countries

Price discrimination Several papers model the welfare implications of price discrimination vs. uniform pricing across countries. Under plausible assumptions about demand dispersion across countries, profit-maximizing price discrimination that leads to prices inversely related to demand elasticities yields higher utilization and improves static efficiency compared to uniform pricing [8]. If R&D is endogenous, price discrimination plausibly also leads to higher R&D investment than uniform pricing [12, 15]. Effects of insurance are not addressed.

Ramsey Pricing Ramsey pricing provides a theoretical foundation for second-best pricing to maximize social welfare, subject to meeting a revenue constraint to cover joint fixed costs [16-18]. Several papers [9-11] apply this theory of Ramsey optimal pricing (ROP) to consider how drug prices should vary across countries to maximize static efficiency, subject to covering the costs of R&D. Because ROP prices vary inversely with demand elasticity, prices would vary be inversely related to per capita income across countries if true demand elasticities (before insurance) vary inversely with per capita income. Under certain conditions with respect to competitive entry and separable markets, the price differentials that an unregulated profit-maximizing manufacturer would set voluntarily would be Ramsey optimal [9,10]. Optimal price differentials might exceed ROP differentials if high income countries have altruistic concerns for low income countries [11]. If a global regulator implements ROP, the effect of insurance in one or more markets on prices in those markets is ambiguous. Insurance makes demand less elastic, leading to higher prices in insured markets, but this may be offset if insurance increases manufacturer revenues such that all prices are reduced to satisfy the ROP revenue constraint [19].

None of these models of price *differentials between* countries consider optimal *absolute* price levels or *within-country* price differentials. In this paper we extend this literature by developing a framework to achieve optimal absolute price levels in each country and optimal differentials between countries. By optimal, we mean that, subject to specified assumptions, including that patents are optimally set, social welfare is maximized in the sense of both static and dynamic efficiency. Our basic framework is designed to deliver “second best” solutions, which we argue may be more realistic and appropriate than models that aim for first best efficiency. Our proposed approach differs between countries with and without comprehensive insurance, because insurance fundamentally changes the challenges to efficiency in pharmaceutical markets and therefore changes appropriate solutions.

III. Value-Based Differential Pricing: A. Countries with Comprehensive Insurance

Optimal ICER Threshold for a Representative Consumer: Exogenous Prices and Technology

Assume that each country can be represented by an individual with income Y that is constant in real terms across time periods. Following Garber-Phelps (GP), period-specific utility of income as viewed from period 0 is $v = U_0(Y)$, before discounting or quality of life adjustment. Income is spent on medical care technologies a and b , with prices w_a and w_b respectively, which for now are assumed to be exogenously set, and on other goods. Consumption of medical care in period 0 can affect probability of survival P and quality of life q in future periods. Future utility is discounted by a factor ρ . The expected benefits of medical care can thus be expressed as the sum of discounted expected future QALYs. Expected utility in period 0 can be written:

$$E_0 = U_0(Y - w_a a - w_b b) + v \sum_i [\rho^i q_i \prod P_j] \quad (1)$$

The optimal utilization of technology a using eq.1 is defined by the first order condition:

$$-w_a U'_0 = v \frac{dQ}{da} \quad (2)$$

In this framework, optimal utilization thus implies equating the marginal utility cost of spending on a in period 0, $w_a U'_0$, to the marginal expected benefit, in terms of utility of future QALYs gained from using a , $v dQ/da$. Rewriting:

$$\frac{w_a}{\left(\frac{dQ}{da} \right)} = \frac{v}{U'_0} \quad (3)$$

The left hand side of eq. 3 is the ICER. The numerator is the incremental cost of using a , which here is just w_a . If use of a affects other medical costs (c_a) due to complementarity or substitution, incremental cost of a would include these incremental costs or cost offsets. The

denominator is the expected QALYs gained from using a . Using this approach, at optimal utilization, the ICER is equal to the ratio of future, period-specific utility v to marginal utility in the base period, or willingness-to-pay for medical care.

Application to a Universal Payer: Endogenous Prices and Utilization

Assume that each country operates a universal insurance system including drugs for all citizens, to provide financial protection and reduce the effect of within-country income dispersion on access to medical care. The analysis can still focus on the representative consumer. Prices charged by manufacturers and technology availability are now endogenous and influenced by the design of the insurance. Reasonable financial protection for patients requires that cost-sharing is modest and capped per patient. Manufacturers therefore face relatively inelastic demand which, in the absence of constraints, would lead to prices above the patent-induced level without insurance (supplier moral hazard). Assume that raising funds is costly, and hence that the payer seeks to achieve second best static and dynamic efficiency.

In this context, eq. 3 implies that the payer can indirectly control prices by setting an ICER threshold K (e.g. £30,000 per QALY) that reflects its citizens' willingness-to-pay for medical care:

$$ICER = \frac{(w_a - w_0) + (c_a - c_0)}{dQ} \leq K \quad (3')$$

If we assume a manufacturer is permitted to price freely, subject to this constraint, it would set its price differential over current treatment at the highest level consistent with still meeting the ICER threshold, given the product's incremental cost offsets and effectiveness gain:

$$w_a = w_0 + dc + KdQ \quad (3'')$$

Thus a new product with no incremental benefit would be constrained to price at w_0 . A more effective product or one that substitutes for other services could be priced higher and still meet the ICER threshold. The payer's ICER threshold acts as an indirect control on price, given the product's incremental effectiveness and cost offsets.

Given the manufacturer's choice of price, the payer can achieve appropriate use by limiting coverage to those patients for whom the product is cost-effective at this price and ICER threshold. In the simplest case of a product for a single indication and uniform effects, all

patients with that condition would be treated. For more complex products, there may be a distribution of effectiveness across patient subgroups with different indications or severities. If the firm chooses a high price, the payer would restrict use to patients whose condition/indication implies an expected health benefit sufficient to meet the ICER threshold, whereas if the firm chooses a lower price, the payer can encourage use by patient subgroups with lower expected benefit and still meet the ICER threshold. Thus given the payer's ICER threshold and a distribution of effectiveness across patients, the firm faces a price-volume trade-off similar to a demand curve except that the payer determines which uses will be reimbursed, given the price set by the manufacturer. If the firm can charge only one price, it would select the price yielding the highest expected profit, given the use that the payer would permit at that price. This outcome would be second best efficient. But if effectiveness differs across indications/subgroups, then both static and dynamic efficiency could be enhanced and, at the limit, would be first best, if the firm could vary prices by indication/subgroup, reflecting the drug's differential effectiveness and the payer could costlessly distinguish and pay the appropriate prices for each indication. Such differential pricing within product may become increasingly important and feasible as drugs become more "personalized" based on patient biomarkers and data systems are improved to provide the necessary information at reasonable administrative cost.

Our approach to value-based pricing is similar to that proposed by Claxton et al. [22], with important differences. First, because our approach is grounded in overall utility maximization, the payer's ICER threshold reflects consumers' willingness-to-pay for health gain and the health care system is assumed to be funded accordingly. By contrast, Claxton et al. take the health care budget as given, and the ICER threshold reflects the opportunity cost of current resource use. Second, our approach permits prices that transfer all surplus to manufacturers for the duration of the patent to achieve dynamic efficiency. This reflects our underlying assumption that patent terms are designed to achieve the optimal trade-off between current consumption and R&D for future consumption. Claxton et al. constrain manufacturers to a single price to retain some surplus for payers/consumers during the term of the patent. Implicitly, this assumes a higher preference for current vs. future consumption for pharmaceuticals compared to other areas of innovation that rely on patents. Third, they focus on a single country whereas our approach addresses efficient pricing and utilization in a global context.

We follow Garber et al. [4] and Lackdawalla and Sood [7] in using insurance to achieve efficiency, but our approach differs in important respects. In our value-based differential pricing (VBDP) approach, manufacturer prices are constrained indirectly, without the payer having to set a price. Given the payer's ICER threshold, a drug's maximum reimbursable price is constrained by its incremental effectiveness and cost savings, relative to current treatments. Constraining prices indirectly, through an ICER threshold, avoids having to assume that the payer can measure the counterfactual, pre-insurance, drug-specific consumer surplus for each drug [5-7]. This is "value-based pricing" in that such prices reward incremental effectiveness and cost savings, evaluated at societal willingness-to-pay for QALYs. Under our approach, in the single price case, price and use reflect second best optimal levels and assigns to the firm the full surplus at this level, to incentivize second best dynamic efficiency.

We assume that eligibility for reimbursement/utilization of costly services is determined by the payer's clinical criteria, rather than attempting to set patients' cost-sharing at marginal cost and letting patients determine use, for three reasons. First, having a payer determine and measure the conceptually appropriate marginal cost is impractical and ignores effects on complementary or substitute medical services. Second, even if payers could set co-payments equal to the appropriate marginal cost, and patients accurately assessed their marginal benefit and chose first best utilization, paying uniform top-up payments for this level of use would not be optimal if raising funds is costly, such that second best principles apply. Third, this approach ignores the caps on patient cost sharing that exist in most public and private insurance plans, to provide financial protection. Patients who use costly and/or chronic medications spend beyond the cost-sharing stop-loss, at which point all care is free and cost-sharing cannot constrain prices or use. We therefore assume that co-payments are set at modest levels to deter excess demand and raise funds, and that payers determine eligibility for reimbursed utilization based on whether the patient's expected benefit meets the ICER threshold. This conforms to actual practice in most national insurance systems, where payers define eligibility for costly technologies and only reimburse for patients with approved indications. Similarly, in the US, most private and public health plans have cost-sharing limits such that nominal co-payments rarely determine use of costly products. Patients without coverage or without stop-loss protection are usually eligible for manufacturers' cost-sharing coupons or patient assistance programs. Thus actual use is

constrained mainly by payers' prior authorization and step therapy protocols that limit reimbursement for expensive drugs to patients that meet clinical criteria [20].

Practical Issues in Defining ICER Thresholds

Patient Heterogeneity If a payer applies a single threshold ICER across all individuals, this ICER may differ from the willingness-to-pay for medical care of some individuals. In a world of perfect information, costless insurance and no altruistic concern for the poor, this would violate Pareto efficiency. Under more realistic assumptions, any insurance imposes some restrictions on individual choice. Presumably, citizens choose to establish national health insurance or join private insurance plans because the expected benefits, in financial protection and improved equity, exceed the utility cost of constrained choice. In such contexts, setting rules in ways that aim to reflect ex ante consumer preferences is likely to be superior to strategies that ignore them. How best to elicit such preferences is an important issue for future research.

Moreover, there are ways of accommodating consumer heterogeneity, even in universal insurance systems with a single payer or competing insurers, to provide for both “voice” and “exit.” For example, in a pluralist system of competing insurers such as the US, different health plans could choose different ICER thresholds, implying different levels of patient access and different drug prices, implemented through rebates as already occurs. Individuals would choose a health plan that best reflects their preferences and willingness-to-pay for health. Such differences could be constrained by minimum universal coverage requirements and subsidies, as proposed by the Affordable Care Act of 2010.

Private and public payers, in single payer or competing payer systems, could also vary ICERs by health condition to address social preferences across conditions. An illustrative example is given by the UK NICE. Although NICE has gone further towards uniformity than most payers by publishing an acceptable range for cost per QALY, exceptions exist. NICE has also introduced an “end of life” ICER threshold, reflecting perceptions that society’s willingness-to-pay for health increases with disease severity. Further, in any insurance plan, patients who are deemed ineligible for coverage of a given product could be permitted to pay out of pocket. Thus the potential welfare losses associated with use of uniform ICERs across patients with

heterogeneous preferences could be minimized, subject to any requirements for universal coverage.

Measuring Costs and Effectiveness Measurement of costs and effectiveness raises many practical challenges, both conceptual (which costs should be included?) and empirical (how to estimate uncertain long term benefits?), and implementation is unlikely to be perfect. Addressing these issues is critical but beyond the scope of this paper.

Price Differences Across Countries

Our framework has implications for determining appropriate price differences across countries to achieve static and dynamic efficiency, at least for countries with universal insurance coverage. In such countries, if each public and private payer defines its ICER threshold unilaterally, based on its citizens' willingness-to-pay for health gain; manufacturers are permitted to price up to that threshold; and payers manage utilisation to reimburse for cost-effective use at that price, then the resulting prices and utilization should be consistent with second best static and dynamic efficiency both within each country and across countries.

To consider how willingness-to-pay – and therefore ICER thresholds and VBDP prices -- might vary with income across countries, assume initially that incremental effectiveness, discount factors and quality of life adjusters are invariant with income and that each country ignores the potential effects of its ICER choice on its prices and the global supply of new drugs. For this simplest case, differentiating the right hand side of eq. 3 with respect to income yields:

$$\frac{d}{dY} \left\{ \frac{v}{U'_0} \right\} = \frac{dICER^*}{dY} = 1 - \frac{E_{U'Y}}{E_{UY}} \quad (4)$$

The term E_{UY} measures the elasticity of utility with respect to income ($dU/dY \cdot Y/U$) and is expected to be positive, $E_{UY} > 0$. $E_{U'Y}$ is the elasticity of marginal utility with respect to income ($dU'/dY \cdot Y/U'$) and is also known as relative risk-aversion. If $E_{U'Y} < 0$, individuals are risk-averse and the optimal ICER rises more than in proportion to income. If $E_{U'Y} = 0$, individuals are risk neutral and the optimal ICER rises in proportion to income. Thus if two countries differ in per capita income but otherwise have similar preferences, our model suggests willingness-to-

pay for medical care and the resulting preferred ICER and level of prices and access will likely be higher in the higher income country but the precise relationship to income cannot be predicted from theory.

Incentives to Free Ride A necessary condition for VBDP to be efficient is that each country's ICER(s) should express its citizens' true preferences, rather than attempting to free ride on other countries' willingness-to-pay for the global joint costs of R&D. In practice, a country that accounts for a small share of global drug sales might recognize that setting an artificially low ICER threshold below its true willingness-to-pay reduces the prices that it pays for drugs with only modest effect on its access to drugs or on global incentives for R&D to develop new drugs.

Such free riding incentives exist in any approach to setting pharmaceutical prices and are not unique to VBDP. Free riding may be mitigated under VBDP to the extent that this approach can be applied to all medical technologies, not just drugs, and hence might be less prone to free riding than alternative approaches targeted only at drug price regulation. Further, choosing an ICER threshold below citizens' true willingness-to-pay may constrain access and utilization as well as prices to suboptimal levels.

In practice, countries with universal insurance increasingly use cost-effectiveness as one tool in their regulatory systems for drugs and some other medical technologies. To the extent that these systems reflect country-specific willingness to pay for medical care, current regulatory structures may approximate the VBDP approach. However, there is also growing use of external price referencing for pharmaceuticals, which may imply strategic behavior rather than unilateral choice. In particular, when a high per capita income country references its prices to those paid by lower income per capita countries, this suggests free riding and suboptimal price differentials, rather than country-specific pricing reflected citizens' true preferences, as required for VBDP and for dynamic efficiency generally.

IV. Self-pay markets

Middle and low income countries (MLICs) that lack comprehensive insurance avoid the distorting effect of insurance on manufacturer prices but also lack the financial protection insurance provides. This affects ability to pay especially for lower income citizens. Assuming that countries lacking universal insurance coverage are self-pay markets, in theory the GP model

should apply directly: patients' self-pay demand reflects willingness-to-pay for expected incremental QALY benefits, provided that patients accurately perceive product quality. If certain other conditions are met – specifically, that: 1. Manufacturers of on-patent products can price-discriminate between countries, and 2. generic markets are price-competitive – then prices and utilization in these self-pay markets should also meet conditions for second best static and dynamic efficiency, consistent with VBDP. Affordability for lower income patients also requires manufacturer ability to segment and price discriminate roughly based on income within the country.

In practice, theory suggests and empirical evidence confirms that these conditions for second best static and dynamic efficiency in self-pay markets are violated in many MLICs. First, highly skewed income distributions create incentives for single-price monopolists to charge prices that are high, relative to average per capita income [1], and price-discrimination between income groups within countries is generally not feasible [2]. Second, generic quality is uncertain in most MLICs because generics are not required to meet regulatory standards of bioequivalence to the originator. Uncertain product quality leads to competition focused on brand, rather than price. Evidence confirms that prices of both on-patent and generic drugs are higher, relative to per capita income, in many MLICs compared to industrialized countries, and price competition is weak [2]. This suggests that achieving VBDP in these self-pay markets would thus require regulatory systems to assure product quality and purchasing mechanisms that facilitate price competition between drugs and possibly price discrimination between market sectors based on income. One such approach for HIV, TB and malaria drugs is described in [2]. Design of mechanisms for a broader range of drugs is an important topic but beyond the scope of this paper.

V. Relation between Value-based Differential Prices and Ramsey Pricing

We have argued that if certain conditions are met -- specifically, that each country with comprehensive insurance chooses its ICER threshold(s) based on preferences of a representative citizen; firms set price(s) in each country to meet its ICER threshold(s); and payers assure access for all patients for whom the drug is cost-effective at these prices – then prices and utilization should approximate second best static and dynamic efficiency locally and globally, at least for

these countries. These prices would vary across countries with (but not necessarily in proportion to) per capita income.

As discussed earlier, Ramsey pricing has been proposed [9, 10] as an alternative theoretical framework for determining optimal pharmaceutical price differentials between countries. The question thus arises whether/how VBDP price differentials differ from ROP differentials?

Ramsey pricing is designed to address the problem: Given an exogenously determined investment that serves multiple users, what prices to these users would maximize social welfare, subject to yielding a normal return on the investment (ROI). The ROP solution sets price differentials above marginal cost to minimize welfare loss from consumption below first best levels, subject to covering the firm's fixed cost. Assuming constant marginal cost and a single price per country, this implies prices vary across countries inversely with price elasticity of demand.

Certain differences between ROP and VBDP are implied by their respective objective functions. ROP prices are theoretically designed to yield a normal ROI on a fixed investment, rather than to transfer the entire surplus from an investment innovation to producers, as required for dynamic efficiency and approximated by our VBDP approach. Moreover, the maximand in the standard ROP set up is one period consumer surplus, with no measure of future utility from current consumption and no explicit consumer budget constraint. By contrast, the maximand in the VBDP approach incorporates the effect of current medical utilization on expected lifetime utility, subject to a lifetime budget constraint. The ROP approach solves for price differentials to achieve a target return on fixed past investment, whereas the VBDP approach sets price levels (and differentials by implication) to transfer whatever surplus results from the innovation, valued at each country's willingness-to-pay, to incentivize future R&D. Thus both VBDP and ROP are designed to achieve second best static efficiency, but only VBDP is also designed to achieve dynamic efficiency.

Inspection of both models suggests that optimal prices will vary across countries with income, but precise proportionality is not necessarily implied by either, and the precise relationship could differ. In the case of ROP, prices vary across markets inversely with E , the uncompensated price elasticity of demand (absolute value). From the Slutsky equation (using

absolute values), the uncompensated price elasticity for good j (E_j) is equal to the compensated price elasticity ϵ_j plus the income elasticity of demand η_j times the income share of j :

$$E_j = \epsilon_j + s_j \eta_j \quad (5)$$

Plausible assumptions are that ϵ_j is invariant with income but η_j increases with income.

Differentiating (5) with respect to income yields:

$$\frac{dE_j}{dY} = \frac{ds_j}{dY} \eta_j + \frac{d\eta_j}{dY} s_j \quad (6)$$

The share s is positive and, for a given price, decreases with income, *cet. par.* The first term on the right is therefore negative. The second term is uncertain. The overall expression is likely to be negative i.e. price elasticity decreases with income, unless the last term is large and positive. Thus under plausible assumptions E varies inversely with income, and therefore ROPs vary directly with income across countries. Comparing eqs. 4 and 6, the cross-national relationship between prices and income is not necessarily the same under ROP and VBDP, due to differences in model specification, including how health enters the objective function of the respective models. In the ROP formulation, income affects prices via the income elasticity of demand for health, whereas in the VBDP model the effect of income depends on the elasticity of marginal utility of non-medical consumption. Obviously, there is no unique *a priori* correct way to incorporate health into utility functions and hence no determinate relationship between optimal drug prices and income across countries. However, both the ROP and VBDP approaches suggest that optimal prices will plausibly increase with income.

In our characterization of self-pay markets we have argued [9,10] that since profit maximizing price discrimination leads to the same *relative* prices (inversely related to demand elasticities) as ROP prices, market incentives should lead unregulated price discriminating firms to set optimal price relativities across markets, provided they can segment markets between and within countries, based on demand elasticities. *Absolute* price levels might be constrained to yield only a normal ROI by competitive entry to achieve a monopolistically competitive equilibrium. Whether actual cross-national price differentials in self-pay MLIC markets

approximate ROP differentials cannot be determined because true demand elasticities and marginal costs are not readily observable. However, observed elasticities of prices with respect to average per capita income are close to zero for a sample of MLICs [2], which seems inconsistent with optimal differentials under either ROP or VBDP prices. As discussed above, characteristics of many MLICs that violate necessary conditions for optimal differentials include: skewed income distributions combined with a lack of within-country price-discrimination; product quality uncertainty for generics and copy products due to lack of enforcement of regulatory quality standards; and competition based on brand, not price [2]. These plausibly explain the divergence of observed elasticities of prices with respect to average per capita income from those predicted by both VBDP and ROP.

Conclusions

An increasing number of countries use a range of regulatory systems for setting pharmaceutical prices, but these regulatory approaches are generally ad hoc and not solidly grounded in economics. We show that for countries with universal insurance, if each country unilaterally and non-strategically sets an ICER threshold based on willingness-to-pay for health gain and permits firms to price up to the ICER, while the payer assures reimbursement for all patients whose expected health gain meets the ICER effectiveness threshold, the resulting prices and utilization would be “valued-based” and yield second best static and dynamic efficiency within and across countries. In other words, the resulting price levels and differentials would be appropriate across countries to achieve second best optimal global incentives for utilization and innovation. If it is feasible, in addition, for payers to implement price differentials between patient subgroups/indications to reflect their differences in incremental effectiveness, then VBDP would approximate first best static and dynamic efficiency. Such value-based prices are likely to vary cross-nationally with per capita income and be broadly consistent with Ramsey Optimal Prices but the precise relationship with per capita income is indeterminate *a priori*.

Countries lacking comprehensive insurance avoid the distorting effect of insurance on manufacturer prices, but also lack insurance’s role in financial protection and assuring affordability of medical care, regardless of income. In such self-pay markets, unregulated price discrimination by firms could in theory lead to Ramsey optimal price differentials, which may approximate second best static and dynamic efficiency, provided certain necessary conditions are

met. These include that consumers are well-informed about product quality and firms can segment and price-discriminate within as well as between countries, to reflect the skewness of income within many MLICs.

To the extent that countries with universal insurance use cost-effectiveness as a criterion for reimbursement and these systems reflect country-specific willingness to pay for medical care, current regulatory structures for pharmaceuticals may approximate the value-based differential pricing we model, thereby making efficient use of existing drugs and sending appropriate signals to the global pharmaceutical industry about future investment in R&D. However, other features of regulatory systems are inconsistent with the necessary conditions for efficiency outlined here. Notably there is considerable use of external price referencing which, by seeking to link prices across countries that do not have similar levels of income per capita and preferences over health and other goods, undermines the condition that each country's ICER(s) express its citizens' true preferences and willingness-to-pay for health gain. In addition, the requirement for first best static and dynamic efficiency – differential pricing by indication and patient subgroup – is currently not feasible for most drugs in most health systems. This may change with the arrival of new biological drugs with multiple indications of different value and increased availability of biomarkers and diagnostics to predict heterogeneity of patient response. If data and information systems also permit differentiated reimbursement, this would permit an ICER-based system to accommodate differential pricing across indications, to accommodate differential value and move towards first best.

Key assumptions for self-pay markets to deliver efficient prices are not present in MLICs, notably drug quality is often uncertain, price competition works poorly and segmentation and price discrimination within countries is often infeasible. These factors contribute to price levels that appear inconsistent with static and dynamic efficiency. Designing systems in MLICs to deliver consistent quality, facilitate price competition and increase affordability while preserving some contribution to and incentives for R&D are important issues for future research.

REFERENCES

- [1] Flynn, S. Hollis, A. Palmedo, M. (2009) An economic justification for open access to essential medicine patents in developing countries. *J. Law, Medicine, Ethics*, 37(2):184-208.
- [2] Danzon, P.M., Mulcahy, A. and Towse, A. (2011). "Pharmaceutical Prices in Emerging Markets: Effects of Income, Competition and Procurement." NBER Working Paper 17174. Cambridge MA.
- [3] Garber, A. and Phelps, C. (1997). Economic foundations of cost-effectiveness analysis. *Journal of Health Economics*, 16, 1-31.
- [4] Garber, A., Jones, C.I. and Romer, P.M. (2006). "Insurance and Incentives for Medical Innovation." *Forum for Health Economics and Policy: Vol.9: Iss.2 (Biomedical Research and the Economy, Article 4*
- [5] Jena, A.B. and Philipson, T.J. (2008) Cost Effectiveness Analysis and Innovation. *Journal of Health Economics*; 27: 1224-1236
- [6] Lakdawalla, D. and Sood, N. (2005) Insurance and Innovation in Health Care Markets. September 2005. NBER Working Paper 11602. Cambridge MA
- [7] Lakdawalla, D. and Sood, N. (2009). Innovation and the welfare effects of public drug insurance. *Journal of Public Economics*; 93: 541-8
- [8] Malueg, D., and Schwartz, M. (1994). "Parallel imports, demand dispersion, and international price discrimination." *J Int'l Econ* 37:167-195.
- [9] Danzon, P. (1997). "Price discrimination for pharmaceuticals: welfare effects in the US and EU." *Int'l J Econ Business* 4(3):301-321.

- [10] Danzon, P. and Towse, A. (2003). "Differential pricing for pharmaceuticals: reconciling access, R&D and patents." *Int J Health Care Finance and Econ* 3:183-205.
- [11] Jack, W. and Lanjouw, J. O. (2005). Financing pharmaceutical innovation: How much should poor countries contribute? *The World Bank Economic Review*, 19(1), 45-67.
- [12] Szymanski, S. and Valletti, T. (2005). Parallel trade, price discrimination, investment and price caps. *Economic Policy*, 20(44), 705-749.
- [13] Ballard, C. L., Shoven J. B. and Whalley, J. 1985 "General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States." *American Economic Review* 75(1): 128-138.
- [14] Goldman, D. and Philipson, T.J. 2007. "Integrated Insurance Design in the Presence of Multiple Medical Technologies," *American Economic Review* 97(2):427-432, May.
- [15] Valletti, T. M. and Szymanski, S. (2006). Parallel trade, international exhaustion and intellectual property rights: A welfare analysis. *The Journal of Industrial Economics*, 54(4), 499-526.
- [16] Ramsey F (1927). A contribution to the theory of taxation. *Economic Journal* 37: 47-61
- [17] Baumol, W. and Bradford, D. (1970). "Optimal departures from marginal cost pricing." *AER* 60(3):365-283.
- [18] Braeutigam, R. (1984). "Socially optimal pricing with rivalry and economies of scale." *RAND J Econ* 15(1):127-134
- [19] Barros, P. and Martinez-Giralt, X. (2008). On international cost-sharing of pharmaceutical R&D. *Int J Health Care Finance and Econ* 8:301-312.

[20] Danzon, P. and Taylor, E. 2010. Drug Pricing and Value in Oncology. *The Oncologist*. 15(1)24-31.

[21] Office of Fair Trading (2007). *The Pharmaceutical Price Regulation Scheme: An OFT market study*. OFT. London

[22] Claxton, K., Briggs, A., Buxton, M.J., Culyer, A.J., McCabe, C., Walker, S. and Sculpher M.J. (2008). Value based pricing for NHS drugs: an opportunity not to be missed? *BMJ*; 336:251-4.