Reference pricing with elastic demand for pharmaceuticals*

Ricardo Gonçalves†
Faculdade de Economia e Gestão, Universidade Católica Portuguesa (Porto)
Vasco Rodrigues
Faculdade de Economia e Gestão, Universidade Católica Portuguesa (Porto)
February 2012

PRELIMINARY DRAFT

Abstract

This paper analyzes pharmaceutical producers’ (of branded and generic drugs) pricing decisions in a context which combines the classical Hotelling framework (commonly adopted in the literature) with the possibility that consumers’ demands are price elastic. We analyze with particular interest two commonly adopted reimbursement mechanisms: fixed percentage reimbursement (FPR) and reference pricing (RP). We find that assuming an elastic demand has a significant impact under a RP mechanism, as it allows firms to raise their prices when the copayment rate increases. Contrary to previous literature, we show that RP may be worse than FPR in terms of social welfare for sufficiently high copayment rates. From a policy viewpoint, our results indicate that the introduction of RP is not necessarily always beneficial (compared to FPR) and thus constitute a model towards which all reimbursement mechanisms should converge.

JEL Classification: I18, L13
Keywords: reference pricing, reimbursement mechanisms, elastic demand

*Financial support from Fundação para a Ciência e Tecnologia (project PTDC/EGE-ECO/100296/2008) is gratefully acknowledged.
†Corresponding author. Postal address: Faculdade de Economia e Gestão, Universidade Católica Portuguesa (Porto), Rua Diogo Botelho, 1327, 4169-005 Porto, Portugal. E-mail: rgoncalves@porto.ucp.pt.
1 Introduction

The extant literature contains a number of strong claims on the relative properties of different pharmaceutical reimbursement mechanisms, and especially on the merits of reference pricing as compared to fixed percentage reimbursement schemes. In this paper, we show that some of those claims are driven by specific modelling assumptions and would, at least, be mitigated under more general assumptions.

In a large number of countries, individuals only pay a fraction of the price of the prescription pharmaceuticals they consume, whilst the government or other third-party payers are responsible for the remainder. Two widely adopted mechanisms to determine the relative contribution of the patient and the government are fixed percentage reimbursement (FPR) and reference pricing (RP). Under FPR, the government pays a fixed percentage of the pharmaceutical’s price and the patient is responsible for the remainder. Thus, the patient’s payment corresponds to a fixed percentage of the price, also known as the copayment rate. By contrast, under RP, within each cluster of pharmaceuticals\(^1\), one is chosen as the ‘reference’ for the determination of the government’s contribution: the government only pays a fixed fraction of the ‘reference’ pharmaceutical’s price, even if the consumer actually buys a more expensive alternative. Naturally, RP is only viable if more than one pharmaceutical is available for the treatment of the same condition and is especially used when low-price generic competition exists.\(^2\)

Reference pricing has attracted a lot of attention in the literature and a large consensus has emerged regarding its merits. For instance, Brekke et al. (2007) show that RP leads to lower prices than FPR mechanisms. They further show that the two mechanisms are equivalent from a social welfare perspective but RP is superior if, adopting a public payer’s perspective, the producers’ profits are excluded from the analysis. Merino-Castelló (2003) also shows that RP enhances price competition and leads to significant reductions in the price of branded pharmaceuticals. On a cautionary note, whilst also acknowledging the potential benefits of this mechanism, Miraldo (2009) notes that if firms anticipate the introduction of RP, they will optimally revise their prices upwards ahead of that introduction, thus significantly limiting its expected effectiveness. Nevertheless, Brekke et al. (2009) and Brekke et al. (2011) present empirical evidence of significant price reductions for all pharmaceuticals following the introduction of RP in Norway.

Most of the models in the literature on reference pricing, namely Brekke et al. (2007), Brekke et al. (2011), Merino-Castelló (2003) and Miraldo (2009), are based on a Hotelling framework: firms sell differentiated versions of a pharmaceutical, with differentiation being represented by their position along a Hotelling interval, and compete by (simultaneously) setting prices; consumers are uniformly distributed along that interval and their position represents their preferences towards

---

\(^1\)Several criteria could be used to cluster pharmaceuticals: chemical, pharmacological or therapeutic (Miraldo, 2009).

\(^2\)For a detailed review of reference pricing, see López-Casasnovas and Puig-Junoy (2000).
the potential varieties of the product; any mismatch between the locations of firms and consumers generates a utility loss to the latter, often described as a ‘transportation cost’, which is added to the price of the pharmaceutical; each consumer buys one unit of the pharmaceutical if total cost - price plus transportation cost - does not exceed his reservation price. Authors commonly assume this restriction to be fulfilled thus implying that every consumer buys one unit of the product. In this type of setting, it is not surprising that social welfare is equal across reimbursement regimes (e.g. as in Brekke et al., 2007), as there are no quantity effects. Reimbursement regimes merely determine who captures the surplus created but not its amount.

This paper’s main purpose is to analyze the impact of reimbursement mechanisms such as FPR and RP in a setting where each consumer’s demand depends on prices. This strikes us as a realistic assumption, at least for some health conditions. Consider, for instance, conditions that result in recurring episodes of symptoms of variable intensity, such as sleeping disorders, muscular pain or migraines. It is entirely plausible that the patient will react to the price of the pharmaceutical he has been prescribed when deciding whether to take it on a specific occasion. Thus, over time, the quantity he buys may be inversely related to price. On the other hand, our setting may be less adequate for pharmaceuticals that are used to treat serious non-recurring conditions, in which case the patient will often be limited to the choice of following his physician’s prescription, and buying the pharmaceutical, or not buying it at all. In this case, the standard Hotelling framework used by Brekke et al. (2007) seems more fitting.

We obtain two novel results. First, we find that, under RP, prices are increasing in the share of the pharmaceutical’s price borne by the consumer, converging to FPR prices as that copayment rate approaches one. This contrasts, for instance, with Brekke et al. (2007) or Gonçalves et al. (2011), who find that, under RP, prices are independent of the copayment rate. Second, we show that, for sufficiently high copayment rates, social welfare under RP is lower than under FPR, a result which runs contrary to the previous literature boasting the merits of RP.

The differences with respect to previous literature’s results are mainly determined by the impact of the elastic demand assumption on the properties of RP mechanisms, as our results for FPR are in line with the previous literature. In a nutshell, whereas under FPR an increase in the copayment rate has no impact on the relative prices faced by the consumer, as all prices rise by the same percentage, under RP this is not so: an increase in the copayment rate changes the relative prices faced by the consumer in favor of the more expensive pharmaceuticals. Ceteris paribus, this would increase their market share but their producers take the opportunity to raise their prices. Prices being strategic complements, other firms respond by also raising prices. Thus, in equilibrium, higher copayment rates lead to higher prices, explaining our first result. Under FPR, social welfare does not depend on the copayment rate: firms’ ‘mark-up’ their prices (and profits) in an inversely proportional manner with respect to the copayment rate such that the amount the consumer has to pay does not change (and, thus, consumer surplus does not depend on the copayment rate). Under
RP, however, higher copayment rates lead to higher prices and lower consumption levels and, thus, to a fall in consumer surplus that is only partially compensated by the increase in profits (net of government expenditure). For sufficiently high copayments rates, this brings social welfare under RP below what it would be with FPR, hence explaining the second result.

From a policy viewpoint, two conclusions emerge: firstly, from a consumer’s perspective, price differences across reimbursement mechanisms are likely to be lower when copayment rates are relatively high. In such a scenario, the introduction of RP is unlikely to result in significant price reductions (and, thus, significant government expenditure savings) compared to FPR. Secondly, when copayment rates are relatively high, RP may lead to lower social welfare levels and its introduction may not be advised. Clearly, and from a broader perspective, our results indicate that the introduction of RP is not necessarily always beneficial and thus does not constitute a model towards which all reimbursement mechanisms should converge.

This paper has the following structure: section 2 describes the base model; section 3 contains our main results and section 4 looks at their welfare implications; finally, section 5 concludes.

2 Model

Our model is loosely inspired by Rath and Zhao (2001), who combine horizontal differentiation, à la Hotelling, with elastic demand for each consumer.

Consumers are assumed to buy a Hicksian composite good (denoted by h) and a differentiated pharmaceutical (denoted by d). This differentiated product exists in two varieties: branded (in which case \( d = b \)) and generic (in which case \( d = g \)). We use the terms ‘incumbent’ and ‘entrant’ to refer to the producers of the branded and generic varieties of the differentiated product, respectively. These two firms compete by simultaneously setting prices.

The consumer's utility from consuming \( q_h \) units of the Hicksian good and \( q_d \) units of the differentiated good is \( u(q_h, q_d) = q_h - 1/q_d - f_d \), where \( f_d \) is a disutility parameter discussed in more detail below.\(^3\) In this utility function, increased consumption of the differentiated product reduces a negative element that can be interpreted as ‘disease’ or ‘discomfort’. The marginal utility associated with consuming the differentiated product is decreasing, i.e. increasing consumption levels contribute gradually less towards the treatment of the underlying health condition.

The consumer’s budget constraint is given by \( p_h q_h + p_d q_d = m \), where \( m \) represents the consumer’s income. When we normalize \( p_h = 1 \), the demand functions that result from utility maximization subject to the budget constraint are given by:

\(^3\)Several utility functions could be used to represent a consumer’s utility in the consumption of these two goods; for instance, Rath and Zhao (2001) assume a utility function which gives rise to linear demand functions. However, many such functions (including that of Rath and Zhao, 2001) make the model untractable and equilibrium prices cannot be explicitly obtained. The exact specification of the utility function (and its implicit demand functions) may not be critical to our results, as we discuss in Section 5.
\[
q_b = m - p_d^{1/2} \\
q_d = \frac{1}{p_d^{1/2}} \tag{1}
\]

Note, in particular, that demand for the differentiated product is isoelastic and the (absolute value of) demand elasticity is equal to 1/2.

We assume that consumers differ in the degree to which they distrust the generic pharmaceutical to be therapeutically equivalent to the branded product. This seems plausible not only among the general public but among trained physicians as well. Kesselheim et al. (2008, p. 2514) note that “...generic drugs are chemically equivalent to their brand-name counterparts in terms of active ingredients but may differ in peripheral features, such as pill color or shape, inert binders and fillers, and the specific manufacturing process”. However, “...[s]ome physicians and patients have expressed concern that bioequivalent generic and brand-name drugs may not be equivalent in their effects on various clinical parameters, including physiological measures such as heart rate or blood pressure, important laboratory measurements, and outcomes such as health system utilization or mortality” (Kesselheim et al., 2008, p. 2515). The scientific debate on whether bioequivalence translates onto treatment effect equivalence is still ongoing, but at the very least suggests that it is plausible and possibly rational for a consumer to “suspect” that a generic drug may not have the same treatment effect as the branded drug.\footnote{Kesselheim et al. (2008) conclude, in their meta-analysis of various studies, that generics and branded drugs have similar clinical outcomes in the treatment of cardiovascular diseases; by contrast, Borgherini (2003) reaches an opposite conclusion - that bioequivalence and therapeutic effectiveness are not necessarily the same - for psychoactive drugs; in line with those findings, Meredith (2003) suggests that average measures of bioequivalence are not necessarily sufficient to ensure treatment effect equivalence for some drugs and notes that “...measures of bioequivalence that take into account variability among subjects and between healthy populations and patient populations may be more appropriate” (p. 2887).}

The parameter \( f_d \) in the utility function measures the utility loss the consumer suffers by buying a drug he trusts less than the branded product. By definition, \( f_b = 0 \), as the branded product is the benchmark against which trust is measured. As for consumers that buy the generic, we assume \( f_g(c) = tc \). The parameter \( c \in [0,1] \) measures the "level of distrust". The parameter \( t > 0 \) measures the negative impact of this distrust on the consumer’s utility. We assume \( c \) to be uniformly distributed in the \([0,1]\) interval.

Our model can be interpreted in two ways. As an analogy with previous models in the RP literature, it is possible to describe the \([0,1]\) interval as a Hotelling line. In this interpretation, the entrant is located at the left end of the interval and uses mill pricing: consumers bear transportation costs. The incumbent, however, uses delivered pricing, making its location not directly relevant to

\footnote{Kobayashi et al. (2011) discusses evidence from various countries (United States, Spain, Portugal, Germany and Norway) where consumers were (or are) skeptical on the effectiveness and safety of generic drugs. In a study on the underlying reasons for generic underuse in Portugal, Quintal and Mendes (2011) find that over 40% of respondents justify it due to their “lack of trust” in generics.}
the determination of the equilibrium prices. It is more intuitive, however, to interpret the model as one of vertical differentiation: in this interpretation, the [0,1] interval is a scale that measures the subjective perceived vertical differentiation between branded pharmaceutical and generic. A consumer with \( c = 0 \) shows "no distrust", i.e. believes the generic to be fully equivalent to the branded pharmaceutical. On the other hand, \( c = 1 \) implies that the consumer has the maximum level of distrust conceivable for a drug that has gone through the licensing process. Thus, if branded and generic prices' were equal, only the consumer with \( c = 0 \) would be indifferent between the two; all others would strictly prefer the branded drug. However, if the generic is cheaper than the branded drug, consumers face a trade-off between their disutility in generic consumption (associated with their distrust of generics) and the higher utility in consumption they can obtain because of lower generic prices.

Further note that we assume this disutility to be independent of the quantity consumed (i.e. lump sum). Stahl (1987), Anderson et al. (1992) and Rath and Zhao (2001) justify this assumption by noting that in standard Hotelling models, where this disutility is a transportation cost associated with the distance between the consumer and the producer’s geographical location, the act of consumption requires a shopping trip, whose cost is largely invariant to the quantity consumed and which translates into economies of scale. In a similar vein, we can regard the disutility associated with purchasing the generic variety as a (fixed) consumer-specific distrust in its treatment effectiveness, which does not depend on the quantity he decides to consume of that drug. Assuming a fixed disutility in the consumption of a generic is equivalent to a decreasing average disutility per unit consumed (somewhat equivalent to the economies of scale mentioned above).\(^6\)

In this context, consumer \( c \in [0,1] \) obtains the following utility from consuming the branded product: \( u_b(q_h,q_b) = q_h - 1/q_h - f_b = q_h - 1/q_b \), because we assume \( f_b = 0 \). Similarly, the utility of consuming the generic product is \( u_g(q_h,q_g) = q_h - 1/q_g - f_g(c) = q_h - 1/q_g - tc \), because we assume \( f_g(c) = tc \). The choice of which variety of the differentiated product (branded or generic) to buy is endogenous: consumers for which \( u_b \geq u_g \) prefer the branded pharmaceutical, whilst all others choose instead to consume the generic. The quantities consumed are given by the demand functions above. Note that the model implicitly assumes that each consumer purchases branded or generic pharmaceuticals, but not both simultaneously: consumers compare the maximum utility they can obtain (given their budget constraint) by consuming (a certain quantity of) each pharmaceutical variety - branded or generic - and choose whichever yields the highest utility.

\(^6\)As an alternative, we could have assumed a constant per unit disutility, i.e. \( f_g(c) = tcq_g \). This would be a sensible assumption if each consumer’s distrust in the generic treatment is likely to be proportional to the quantity consumed, thus implying that, all else equal, consumers would be less confident in the generic’s treatment effectiveness the more they consumed it. Note that empirical studies, such as that of Quintal and Mendes (2011), point to an increased probability of consumers’ acceptance of generic substitution if they have consumed generics before, thus lending some support to our not using this alternative assumption. Nevertheless, when solving the model under this assumption, the results become computationally untractable and no closed form solutions can be found. However, some simulations indicate that the results of Proposition 1 and 2 would hold, but that those of Proposition 3 would not: under this alternative assumption, RP would always lead to higher social welfare levels than FPR.
Both the choice between the generic and branded product as well as the quantities consumed depend on the prices charged to the consumer and these depend on the existence and properties of the reimbursement mechanism. We analyze two reimbursement schemes: a fixed percentage reimbursement (FPR) mechanism (henceforth represented by the superscript ‘F’), in which consumers pay the same percentage (copayment rate), \( \theta \in [0, 1] \), of their desired product’s headline price (the headline price being the price set by the producer); and a reference pricing (RP) mechanism (henceforth represented by the superscript ‘R’), where the consumer’s copayment is based on the price of the ‘reference’ pharmaceutical - typically the generic because it is cheaper and hence leads to a lower reimbursement expenditure by the third-party payer. Under a RP mechanism, the consumer must pay a percentage (copayment rate) of the reference pharmaceutical’s price if he chooses to purchase it, but if he chooses to purchase other pharmaceuticals in a predetermined reference group he must also pay the full price difference between that pharmaceutical’s price and the reference price. In order to highlight the differences between these two reimbursement mechanisms, under a FPR mechanism, effective prices - by which we refer to the price paid by the consumer - are given by:

\[
\hat{p}_k^F = p_k^F - (1 - \theta) p_k^F = \theta p_k^F, \quad k = b, g
\]  

(2)

We define \( p_k^F (k = b, g) \) to be the headline prices set by firms when the FPR mechanism is in place. By contrast, under a reference pricing mechanism, the effective prices are given by (we assume that the reference pharmaceutical is the generic and that the copayment rate, \( \theta \), is similar across mechanisms):

\[
\hat{p}_b^R = p_b^R - (1 - \theta) p_g^R = \theta p_g^R
\]

\[
\hat{p}_g^R = p_g^R - (1 - \theta) p_g^R = \theta p_g^R
\]  

(3)

where \( p_k^R (k = b, g) \) are the headline prices set by firms when the RP mechanism is in place.

We will also refer to the no reimbursement (NR) scenario, whereby patients must support the headline prices in full; this is equivalent to setting \( \theta = 1 \) in equation (2) or (3). Finally, production costs are assumed to be zero and the total number of consumers is set equal to 1, i.e. product quantities can be interpreted as market shares.

3 Pricing

3.1 Fixed Percentage Reimbursement

Under a fixed percentage reimbursement mechanism, where the effective prices for the branded and generic drugs are \( \hat{p}_b^F = \theta p_b^F \) and \( \hat{p}_g^F = \theta p_g^F \) respectively, the “marginal consumer”, \( c_{b,g}^F \), who is indifferent between buying the branded (b) or generic (g) pharmaceuticals is found by solving \( u_b = u_g \). This yields:
As the total number of consumers is assumed to be equal to one, \( c_{b,g} \) gives us the market share of the entrant (which produces the generic drug) and \( 1 - c_{b,g} \) is the market share of the incumbent (which produces the branded drug). Each consumer in the \([0, c_{b,g}]\) interval chooses to purchase \( q_g^F = 1/ (p_g^F)^{1/2} \) units of the generic drug, whilst consumers in the \([c_{b,g}, 1]\) interval choose to purchase \( q_b^F = 1/ (p_b^F)^{1/2} \) units of the branded drug. Let \( \mathbf{p}^F = (p_b^F, p_g^F) \) represent the price vector. Using equation (2), the demand functions are thus given by:

\[
B(\mathbf{p}^F) = q_b^F (1 - c_{b,g}) = \frac{1}{\theta^{1/2} (p_b^F)^{1/2}} \left( 1 - \frac{2\theta^{1/2} \left[ (p_b^F)^{1/2} - (p_g^F)^{1/2} \right]}{t} \right)
\]

\[
G(\mathbf{p}^F) = q_g^F c_{b,g} = \frac{1}{\theta^{1/2} (p_g^F)^{1/2}} \frac{2\theta^{1/2} \left[ (p_b^F)^{1/2} - (p_g^F)^{1/2} \right]}{t}
\]

The profit functions are given by \( \Pi_b(\mathbf{p}^F) = p_b^F \times B(\mathbf{p}^F) \) and \( \Pi_g(\mathbf{p}^F) = p_g^F \times G(\mathbf{p}^F) \) for the incumbent and entrant respectively. Maximizing these functions with respect to \( p_b^F \) and \( p_g^F \) respectively we obtain the best-response functions:

\[
p_b^F = \frac{t^2}{16\theta} + \frac{t}{4\theta^{1/2}} (p_g^F)^{1/2} + \frac{1}{4} p_g^F
\]

\[
p_g^F = \frac{1}{4} p_b^F
\]

In a Nash equilibrium, when we insert the best response function \( p_g^F = \frac{1}{4} p_b^F \) in the first equation of (6), the equilibrium value of \( p_b^F \) is the solution of this non-linear equation:

\[
\frac{15\theta p_b^F - 2t\theta^{1/2} (p_b^F)^{1/2} - t^2}{16\theta} = 0
\]

This equation’s solution is necessarily an approximate root, which leads to the following Nash equilibrium prices:

\[
p_b^F = \frac{4}{36\theta} t^2
\]

\[
p_g^F = \frac{1}{36\theta} t^2
\]
\begin{align}
q_b^F &= \frac{3}{t} \\
q_g^F &= \frac{6}{t}
\end{align} \tag{9}

whilst overall quantities sold are given by:

\begin{align}
B^F &= \frac{2}{t} \\
G^F &= \frac{2}{t}
\end{align} \tag{10}

At the equilibrium prices, \( c_{b,g}^F = 1/3 \), which implies that whilst each individual consumer purchases more units of the generic drug than of the branded drug, the latter is sold to a larger number of consumers. These two effects cancel each other out and in equilibrium each producer sells exactly the same quantity: \( 2/t \) units.

In equilibrium, profits are:

\begin{align}
\Pi_b^F &= \frac{4t}{18\theta} \\
\Pi_g^F &= \frac{t}{18\theta}
\end{align} \tag{11}

When no reimbursement (NR) mechanism exists, \( \theta = 1 \) and we obtain results which are in line with those of Rodrigues et al. (2011), Gonçalves et al. (2011) and Brekke et al. (2007). Note that \( p_k^F = \frac{1}{b} p_k^{NR} \), with \( k = b,g \). Under a fixed percentage reimbursement scheme, firms mark-up their headline prices in an inversely proportional manner relative to the copayment rate: the higher is the copayment rate, the lower is that mark-up under a FPR mechanism. This implies that the effective price paid by consumers is equal in both cases and therefore equilibrium quantities are also the same, i.e. \( \bar{p}_k^F = \bar{p}_k^{NR} \) \((k = b,g)\), \( B^F = B^{NR} \) and \( G^F = G^{NR} \). Profit levels become \( \Pi_j^F = \frac{1}{b} \Pi_j^{NR}, j = b,g, \) i.e. firms also increase their profits in an inversely proportional manner relative to the copayment rate. Headline prices, \( p_b \) and \( p_g \), are lowest when the patient’s copayment rate is maximal, i.e. when \( \theta = 1 \).

Consumer surplus (CS) is given by the sum of the surplus of buying the branded and generic pharmaceuticals (in equilibrium, \( c_{b,g}^F = 1/3 \)):

\[
CS^F = \int_0^{1/3} (q_h^F - 1/q_g^F - tc) \, dc + \int_{1/3}^1 (q_h^F - 1/q_b^F) \, dc \\
= m - \frac{11t}{18}
\] \tag{12}
Note that CS does not depend on $\theta$; as outlined above, the existence of a copayment rate has no impact on the effective prices consumers must face, $\tilde{p}_b$ and $\tilde{p}_g$. Hence, the existence of a FPR mechanism (through parameter $\theta$) does not affect consumer surplus, i.e. $CS^F = CS^{NR}$.

Total profits, which are equivalent to total pharmaceutical expenditure, are given by:

$$\Pi^F = \Pi_b^F + \Pi_g^F = \frac{5t}{18\theta}$$

Government (or other third-party payers) expenditure with pharmaceuticals is a proportion $(1 - \theta)$ of total pharmaceutical expenditure:

$$G^F = (1 - \theta) \Pi^F = (1 - \theta) \frac{5t}{18\theta}$$

Note that government expenditure is decreasing with $\theta$. This result has two causes: (i) first, there is the direct effect that as $\theta$ increases, the reimbursement rate $(1 - \theta)$, applicable to total pharmaceutical expenditure, decreases; on the other hand, there is the strategic effect that producers respond to higher copayment rates by reducing headline prices, thus bringing down total pharmaceutical expenditure.

### 3.2 Reference Pricing

Under a reference pricing mechanism, the effective prices for the branded and generic drugs are $\tilde{p}_b^R = p_b^R - (1 - \theta) p_g^R$ and $\tilde{p}_g^R = \theta p_g^R$ respectively (see equation (3)) and the “marginal consumer”, $c_{b,g}^R$, who is indifferent between buying the branded ($b$) or generic ($g$) pharmaceuticals is again found by solving $u_b = u_g$, which in this case yields:

$$c_{b,g}^R = \frac{2 (\tilde{p}_b^R)^{1/2} - 2 (\tilde{p}_g^R)^{1/2}}{t} = \frac{2 [p_b^R - (1 - \theta) p_g^R]^{1/2} - 2 \theta^{1/2} (p_g^R)^{1/2}}{t}$$

Similarly to the FPR case, $c_{b,g}^R$ gives us the market share of the entrant (which produces the generic drug) under a RP mechanism. Each consumer in the $[0, c_{b,g}^R]$ interval chooses to purchase $q_g^R = 1/ (\tilde{p}_g^R)^{1/2}$ units of the generic drug, whilst consumers in the $[c_{b,g}^R, 1]$ interval choose to purchase $q_b^R = 1/ (\tilde{p}_b^R)^{1/2}$ units of the branded drug. Let $p^R = (p_b^R, p_g^R)$ represent the price vector. The demand functions are thus given by $B(p^R) = q_b^R (1 - c_{b,g}^R)$ for the branded drug and by $G(p^R) = q_g^R c_{b,g}^R$ for the generic drug. The profit functions are given by $\Pi_b(p^R) = p_b^R \times B(p^R)$ and $\Pi_g(p^R) = p_g^R \times G(p^R)$. When we maximize the latter profit function with respect to $p_g^R$, the first-order condition for the entrant (after some simplifying steps) yields:

$$\partial \Pi_g(p^R) / \partial p_g^R = 0 \Leftrightarrow p_b^R - 2 (1 - \theta) p_g^R - 2 \theta^{1/2} (p_g^R)^{1/2} (p_b^R - (1 - \theta) p_g^R)^{1/2} = 0$$

This non-linear equation has two solutions, one of which results in $p_b^R < p_g^R$ whenever $\theta > 1/4$. This violates the assumption that the generic drug is the reference pharmaceutical because it is
cheaper, and hence the best-response function of the entrant is assumed to correspond to this equation’s other solution:

\[ p^R_g = \frac{1 - \theta^{1/2}}{2 (1 - \theta) p^R_b} \]  

(17)

Maximizing the incumbent’s profit function with respect to \( p^R_b \) and setting it equal to zero \( (\partial \Pi_b (p^R) / \partial p^R_b = 0) \) results in the following first-order condition:

\[
\frac{tp^R_b - 2t (1 - \theta) p^R_b - 4 \left[ p^R_b - (1 - \theta) p^R_g \right]^{3/2} + 2\theta^{1/2} p^R_b (p^R_g)^{1/2} - 4\theta^{1/2} (1 - \theta) (p^R_g)^{3/2}}{2t \left[ p^R_b - (1 - \theta) p^R_g \right]^{3/2}} = 0
\]  

(18)

Substituting the entrant’s best-response function (equation (17)) into this equation, we obtain the incumbent’s Nash equilibrium price and subsequently substituting it into equation (17), we also obtain the entrant’s equilibrium price:

\[
p^R_b = \frac{2\theta \left( 1 + \theta^{1/2} \right)}{4 \left( 1 + 2\theta^{1/2} \right)^2 t^2}
\]

\[
p^R_g = \frac{\theta}{4 \left( 1 + 2\theta^{1/2} \right)^2 t^2}
\]  

(19)

Under a RP mechanism, the equilibrium individually consumed quantities are given by:

\[
q^R_b = \frac{2 \left( 1 + 2\theta^{1/2} \right)}{\left( \theta + \theta^{1/2} \right) t}
\]

\[
q^R_g = \frac{2 \left( 1 + 2\theta^{1/2} \right)}{\theta t}
\]  

(20)

whilst overall quantities sold are given by:

\[
B^R = \frac{2}{\theta^{1/2} t}
\]

\[
G^R = \frac{2}{\theta^{1/2} t}
\]  

(21)

At the equilibrium prices, \( c^R_{b,g} = \theta^{1/2} / \left( 1 + 2\theta^{1/2} \right) \leq 1/3, \forall \theta \in [0, 1] \) (in particular, \( c^R_{b,g} \) is an increasing function of \( \theta \)) which implies that, as under the FPR mechanism, whilst each individual consumer purchases more units of the generic drug than of the branded drug, the latter is sold to a larger number of consumers. However, also as under the FPR mechanism, these two effects cancel each other out and each producer sells exactly the same quantity: \( 2 / \left( \theta^{1/2} t \right) \) units.

It is particularly relevant to highlight the following (novel) result:
Proposition 1 Under a RP mechanism, both headline \( (p^R = (p^R_b, p^R_g)) \) and effective equilibrium prices \( (\hat{p}^R = (\hat{p}^R_b, \hat{p}^R_g)) \) are increasing in \( \theta \).

Proof. In equilibrium, under a RP mechanism, headline prices are given by equation (19), whose derivative with respect to \( \theta \) is:

\[
\frac{\partial p^R_b}{\partial \theta} = \frac{3 \theta^{1/2} + 2 \theta + 2}{4 \left(1 + 2 \theta^{1/2}\right)^3} t^2
\]

\[
\frac{\partial p^R_g}{\partial \theta} = \frac{t^2}{4 \left(1 + 2 \theta^{1/2}\right)^3}
\] (22)

Both derivatives are positive for any \( t > 0 \) and \( \theta \in [0,1] \). Effective equilibrium prices are easily calculated using equations (3) and (19), yielding \( \hat{p}^R_b = \theta \left(1 + \theta^{1/2}\right)^2 \frac{t^2}{4 \left(1 + 2 \theta^{1/2}\right)^2} \) and \( \hat{p}^R_g = \theta^2 t^2 \left(1 + 2 \theta^{1/2}\right)^2 \). The derivative of effective prices with respect to \( \theta \) is:

\[
\frac{\partial \hat{p}^R_b}{\partial \theta} = \frac{1 + 3 \theta^{1/2} + 4 \theta + 2 \theta^{3/2}}{4 \left(1 + 2 \theta^{1/2}\right)^3} t^2
\]

\[
\frac{\partial \hat{p}^R_g}{\partial \theta} = \frac{\theta + \theta^{3/2}}{2 \left(1 + 2 \theta^{1/2}\right)^3} t^2
\] (23)

Again, both derivatives are positive for any \( t > 0 \) and \( \theta \in [0,1] \).

When we compare these equilibrium prices with those obtained under a FPR mechanism, we also obtain the following result:

Proposition 2 Under a RP mechanism, both headline \( (p^R = (p^R_b, p^R_g)) \) and effective equilibrium prices \( (\hat{p}^R = (\hat{p}^R_b, \hat{p}^R_g)) \) are never higher than under a FPR mechanism.

Proof. Recall that headline equilibrium prices under a FPR mechanism are those given by equation (8), whilst under a RP mechanism they are given by equation (19). Using those expressions, it is straightforward to show that \( p^F_b > p^R_b \) for any \( \theta \in [0,1] \) and \( p^F_b = p^R_b \) when \( \theta = 1 \). Similarly, \( p^F_g > p^R_g \) for any \( \theta \in [0,1] \) and \( p^F_g = p^R_g \) when \( \theta = 1 \).

Effective equilibrium prices under FPR are easily calculated using equations (2) and (8): \( \hat{p}^F_b = 4t^2/9 \) and \( \hat{p}^F_g = t^2/9 \). Effective equilibrium prices under RP are given above (see proof of Proposition 1). It is now straightforward to show that \( \hat{p}^R_b > \hat{p}^F_b \) for any \( \theta \in [0,1] \) and \( \hat{p}^R_b = \hat{p}^F_b \) when \( \theta = 1 \). Similarly, \( \hat{p}^R_g > \hat{p}^F_g \) for any \( \theta \in [0,1] \) and \( \hat{p}^R_g = \hat{p}^F_g \) when \( \theta = 1 \).

In order to understand the rationale underlying these two results, Figure 1 contains a graphical representation of equilibrium headline (left) and effective prices (right) under the FPR and
RP mechanisms, assuming $t = 0.5$. First, note that headline prices under a FPR mechanism are decreasing with $\theta$ whilst under a RP mechanism they are increasing (Proposition 1). Under a FPR mechanism, an increase in the copayment rate ($\theta$) would lead, ceteris paribus, to an increase in effective prices, therefore reducing the quantity consumed of both the branded and generic pharmaceuticals. In this context, both firms find it profit-maximizing to reduce headline prices and, thus, to reduce the (negative) impact on demand. It turns out that, in equilibrium, this reduction is sufficient for effective prices not to change and, thus, for individually consumed quantities and overall sold quantities of both goods to remain constant. Inevitably, this hurts both firms’ profits. As mentioned earlier, this result is similar to that of Brekke et al. (2007). An increase in the copayment rate leads to a reduction in total pharmaceutical expenditure (which is equal to profits) and, consequently, to a reduction in government expenditure, but consumer surplus remains unchanged. From a different point of view, reimbursement through a FPR mechanism can be seen as a public subsidy received by firms; as the copayment rate increases (i.e. as the reimbursement rate decreases), this public subsidy decreases (and becomes equal to zero when $\theta = 1$).

Under a RP mechanism, an increase in $\theta$ (ceteris paribus) would lead to an increase in effective prices by the same amount. However, because it is lower, the effective price of the generic would increase by more than the effective price of the branded pharmaceutical in percentage terms. In comparative terms, the branded product would become relatively less expensive, thus attracting some of the consumers that were previously buying the generic. In this context, the incumbent finds it profit-maximizing to respond to the increase in the copayment rate with an increase in its headline price. As prices are strategic complements, the entrant does the same. In equilibrium, the increase in the copayment rate coupled with the increase in headline prices unequivocally increases effective prices and reduces individually consumed quantities as well as overall demand for both pharmaceuticals.

Note that, as shown in Proposition 2, headline and effective prices are higher under a FPR mechanism (except for $\theta = 1$, in which case they are equal). This suggests that RP is more effective in bringing about price competition than FPR, to the benefit of consumers. This result is in line with that of the existing literature (Brekke et al., 2007; Merino-Castelló, 2003).

Under a RP mechanism, in equilibrium, profits are:

$$
\Pi^R_b = \frac{\theta^{1/2} \left(1 + \theta^{1/2}\right) t}{\left(1 + 2\theta^{1/2}\right)^2}
$$

$$
\Pi^R_g = \frac{\theta^{1/2} t}{2 \left(1 + 2\theta^{1/2}\right)^2}
$$

Note that the result in Proposition 2 does not depend on the value of $t$. This assumption is made here merely with the objective of obtaining a graphical representation of the results and understanding their rationale.
Figure 1: Headline (left) and effective (right) equilibrium prices under FPR and RP (assuming $t = 0.5$)

**Corollary 1** The incumbent’s profits are always increasing with $\theta$, whilst the entrant’s profits are only increasing with $\theta$ when $\theta < 1/4$.

**Proof.** For the incumbent, $\partial \Pi_b^R / \partial \theta = \frac{\theta^{1/2}(1+\theta^{1/2})^2}{2(1+2\theta^{1/2})^2(\theta+\theta^{1/2})^2} > 0$, $\forall \theta \in [0,1]$. For the entrant, $\partial \Pi_g^R / \partial \theta = \frac{t(1-2\theta^{1/2})}{4\theta^{1/2}(1+2\theta^{1/2})^3}$; this is positive only when $\theta < 1/4$. ■

As argued above, an increase in $\theta$ works to the incumbent’s advantage, as it allows it to raise headline prices and profits. By contrast, the entrant’s profits are only increasing with $\theta$ when the copayment rate is relatively low ($\theta < 1/4$); above that threshold, the entrant’s profits are decreasing with $\theta$, as it finds the net effect of higher headline and effective prices and lower demand to be negative.

Consumer surplus (CS) is given by the sum of the surplus of buying the branded and generic pharmaceuticals (in equilibrium, $c_{b,g}^R = \theta^{1/2} / \left(1 + 2\theta^{1/2}\right)$):

$$CS^R = \int_0^{\theta^{1/2}} \left(q_h - 1/q_g - tc\right) dc + \int_{\theta^{1/2}}^1 \left(q_h - 1/q_g\right) dc$$

$$= m - \frac{(5\theta + 2\theta^{1/2} + 4\theta^{3/2})t}{2 \left(1 + 2\theta^{1/2}\right)^2}$$  \hspace{1cm} (25)

Note that consumer surplus under a RP mechanism is decreasing with $\theta$. This is to be expected
because effective prices are increasing with $\theta$ (Proposition 1). Total profits, which are equivalent to total pharmaceutical expenditure, are given by:

$$\Pi^R = \Pi^R_b + \Pi^R_g = \frac{3\theta^{1/2} + 2\theta}{2 \left(1 + 2\theta^{1/2}\right)^2} t$$

(26)

Thus, overall profits are an increasing function of $\theta$: as $\theta$ increases, the increase in profits of the incumbent is sufficient to offset any reduction in profits by the entrant. Government (or other third-party payers) expenditure with pharmaceuticals is a proportion $(1 - \theta)$ of total pharmaceutical expenditure:

$$G^R = (1 - \theta) \Pi^R = (1 - \theta) \frac{3\theta^{1/2} + 2\theta}{2 \left(1 + 2\theta^{1/2}\right)^2} t$$

(27)

Corollary 2. Below a certain threshold $(\theta' \approx 0.14)$, government expenditure is increasing with $\theta$.

Proof. Note that $\partial G^R/\partial \theta = \frac{t(3 - 2\theta^{1/2} - 9\theta - 14\theta^{3/2} - 8\theta^2)}{4\theta^{1/2}(1 + 2\theta^{1/2})^3}$. This derivative is positive whenever $\theta$ is below the threshold $\theta' \approx 0.14$.

For low values of $\theta$, government expenditure is increasing with $\theta$: producers respond to higher copayment rates by increasing prices but, at these low levels of copayment, this has little impact on demand; resulting price rises lead to a higher pharmaceutical expenditure which implies higher reimbursements. Above a certain threshold $(\theta' \approx 0.14)$, this no longer is the case: the reduced reimbursement rate applied to higher headline prices and lower total demand decreases government expenditure.

4 Welfare analysis

We now turn our attention to the welfare implications of our results. We first focus on social welfare (SW) - the sum of consumer and producer surplus minus government expenditure with pharmaceuticals - and show that:

Proposition 3. There exists a threshold value ($\tilde{\theta} = 2/3$) below which social welfare is higher under a RP mechanism and above which social welfare is higher (or equal) under a FPR mechanism.

Proof. Under a FPR mechanism, social welfare is given by $SW^F = CS^F + \Pi^F - G^F$. Using the expressions in equations (12), (13) and (14) we obtain:

$$SW^F = m - \frac{t}{3}$$

(28)
By contrast, under a RP mechanism, social welfare is given by \( SW^R = CS^R + II^R - G^R \). Using the expressions in equations (25), (26) and (27) we obtain:

\[
SW^R = m - \frac{\left(2\theta^{1/2} + 5\theta + \theta^{3/2} - 2\theta^2\right) t}{2 \left(1 + 2\theta^{1/2}\right)^2} \tag{29}
\]

The difference between social welfare under the two reimbursement mechanisms is given by:

\[
SW^F - SW^R = \frac{\left(-2\theta^{1/2} + 7\theta + 3\theta^{3/2} - 6\theta^2 - 2\right) t}{6 \left(1 + 2\theta^{1/2}\right)^2} \tag{30}
\]

The denominator is always positive and \( t \) is also assumed to be strictly positive. Therefore, the sign of this expression depends on the numerator, which is a concave non-linear function. It has two roots: \( \hat{\theta} = 2/3 \) and \( \theta = 1 \). Therefore, when \( 0 \leq \theta < \hat{\theta} = 2/3 \), we have \( SW^F - SW^R < 0 \); by contrast, when \( 1 \geq \theta > \hat{\theta} = 2/3 \), we have \( SW^F - SW^R \geq 0 \) (above the threshold \( \hat{\theta} = 2/3 \), \( SW^F - SW^R > 0 \) for all values of \( \theta \) strictly below 1 and \( SW^F - SW^R = 0 \) when \( \theta = 1 \), as this is the second root of the non-linear function).

Brekke et al. (2007) suggest a different welfare measure which may be more relevant to countries which do not have a significant pharmaceutical industry. In those cases, it is unlikely that profits would enter the welfare function. Thus, welfare \( (W) \) can be more narrowly defined in a public payer perspective, only taking into account consumer surplus net of government expenditure. Following this approach, we can state that:

**Proposition 4** From a public payer’s welfare perspective, RP is never worse than FPR, i.e. \( W^R \geq W^F \), \( \forall \theta \in [0, 1] \).

**Proof.** Using the expressions in equations (12) and (14), welfare in a public payer’s perspective under the FPR mechanism is given by:

\[
WF = CS^F - G^F = m - \frac{t}{3} - \frac{5t}{18\theta} \tag{31}
\]

whilst under a RP mechanism it is given by (see equations (25) and (27)):

\[
WR = CS^R - G^R = m + \frac{\left(-5\theta^{1/2} - 7\theta - \theta^{3/2} + 2\theta^2\right) t}{2 \left(1 + 2\theta^{1/2}\right)^2} \tag{32}
\]

The difference in welfare under the two reimbursement mechanisms is given by:

\[
WF - WR = \frac{\left(-20\theta^{1/2} - 26\theta + 21\theta^{3/2} + 39\theta^2 + 9\theta^{5/2} - 18\theta^3 - 5\right) t}{18\theta \left(1 + 2\theta^{1/2}\right)^2} \tag{33}
\]
The denominator is always positive and \( t \) is also assumed to be strictly positive. Therefore, the sign of this expression depends on the sign of the concave non-linear function in the numerator, which is negative for \( \forall \theta \in [0, 1) \) (in which case \( W^F < W^R \)) and equal to zero when \( \theta = 1 \) (in which case \( W^F = W^R \)).

Figures 2 contains a graphical representation of consumer surplus, profits net of government expenditure, social welfare and welfare from a public payer’s perspective under both reimbursement regimes, assuming \( t = 0.5 \) and \( m = 1 \).\(^8\) First, note that social welfare under a FPR mechanism does not depend on \( \theta \); this result is similar to Brekke et al. (2007) (Proposition 6). As argued earlier, under a FPR mechanism, pharmaceutical firms effectively receive a public subsidy but consumers do not benefit from lower (effective) prices: the mark-up of headline prices (proportional to the reimbursement rate) is fully borne by the government and hence this transfer (effectively a public subsidy) is neutral from a social welfare perspective. From a policy viewpoint, this result clearly indicates the need for pharmaceutical price regulation prior to inclusion in a FPR mechanism.

By contrast, under reference pricing, social welfare depends on \( \theta \) : as it increases, headline and effective prices also increase, thus reducing consumer surplus, increasing profits and (above a certain threshold) reducing government expenditure. The net effect of \( \theta \) on social welfare depends on the relative contributions of these offsetting effects: the negative effect on consumer surplus and the positive effect on profits and government expenditure.\(^9\) For a large range of values for \( \theta \) \((\theta \in [0, 0.82])\), with the upper bound being an approximate value), the negative effects dominates and thus social welfare is decreasing with \( \theta \), whilst for higher values of \( \theta \) social welfare is increasing with \( \theta \).

Combining these two results leads to the conclusion (Proposition 3) that a threshold exists below which RP yields a higher social welfare but above which it is FPR which yields a higher social welfare. This is not totally surprising: as \( \theta \) increases, consumer surplus under RP decreases and approaches that of FPR (which is constant); by contrast, profits net of government expenditure increase and approach those of FPR (see Figure 2). Above that threshold, the (marginally negative) contribution that profits net of government expenditure make towards social welfare under RP (relative to FPR) becomes more significant than the (marginally positive) contribution of consumer surplus.

When profits are not considered in the welfare measure - as in Proposition 4 - then only consumer surplus and government expenditure should be considered. Under a RP mechanism, consumer surplus is decreasing with \( \theta \) but always larger than (or equal to, when \( \theta = 1 \)) under FPR; government expenditure is always decreasing with \( \theta \) under FPR and almost always (above \( \theta \simeq 0.14 \)) decreasing.

\(^8\)Again, note that the results in Propositions 3 and 4 do not depend on the value of \( t \) and \( m \). This assumption is made here merely with the objective of obtaining a graphical representation of the results and understanding their rationale.

\(^9\)Note that, taken together, the sum of profits net of government expenditure is an increasing function of \( \theta \), as we can see in Figure 2.
Figure 2: Consumer surplus and profits net of government expenditure (left), social welfare (middle) and welfare from a public payer’s perspective (right) under FPR and RP (assuming $t=0.5$ and $m=1$)
with \( \theta \) under RP, but under RP always lower than (or equal to, when \( \theta = 1 \)) under FPR. Therefore, from a public payer’s perspective, RP unequivocally leads to higher welfare levels than FPR, and this is particularly so the lower is the copayment rate (or the higher is the reimbursement rate) - when consumer surplus under RP is highest (relative to FPR) and when government expenditure under FPR is also highest (relative to RP).

5 Discussion and conclusion

This paper studies the competitive properties of two commonly adopted reimbursement mechanisms, fixed percentage reimbursement (FPR) and reference pricing (RP), in a setting in which consumers’ individual demand is price-elastic. Contrary to previous literature, we show that RP may be worse than FPR, from a social welfare viewpoint, for sufficiently high copayment rates. This is because, in such a setting, RP has the counterintuitive property that higher copayment rates lead to higher prices. Policy-wise, our results throw some caution in the prevailing idea that introducing RP is likely to generate substantial benefits to consumers (in the form of lower prices) and will necessarily lead to higher social welfare levels. Thus, whereas we show that RP does have some advantages compared to FPR, it is unlikely to be an ‘always better’ reimbursement mechanism.

An important question is whether our results would hold for demand functions different from the very specific one we consider. The main driver of our results is the fact that, under RP, an increase in \( \theta \) leads to a higher percentage increase in the effective price of the generic and, thus, to a change in relative prices in favor of the branded pharmaceutical. This tips market shares in favor of the more expensive product. The key question, then, is whether this is sufficient to offset the direct negative impact on demand for that product that results from the higher copayment rate, creating the incentive for a price increase. Whilst we have not explicitly solved the model under alternative setups, it appears plausible that, for a large family of demand functions, it will. Exploring the conditions in which this happens is a possible avenue for future research.

Clearly, this paper’s results on pricing are amenable to empirical testing, as that which was carried out by Brekke et al. (2009) and Brekke et al. (2011). It is particularly relevant to highlight that our predictions only differ from those of Brekke et al. (2009) and Brekke et al. (2011) in what RP is concerned. In addition, it may be worthwhile to extend the analysis to a third possible, and to the best of our knowledge yet under-researched, reimbursement mechanism: asymmetric fixed percentage reimbursement, through which the reimbursement rates are different for different types of pharmaceuticals, with generics typically attracting higher reimbursement (lower copayment) rates. This type of mechanism has been used, for instance, in Portugal in order to help increase generic penetration at an early development stage. Whilst this type of mechanism introduces more complexity in the analysis, it may also have the merit to uncover more intricate details of firms’ pricing incentives. These are also possible next steps in our research.
References


