CCHE CCES
Canadian Centre for Health Economics
Centre canadien en économie de la santé

Working Paper Series
Document de travail de la série

HOW DOES INSURANCE AFFECT THE PRICE OF DRUGS: A GRAPHICAL ANALYSIS

Audrey Laporte, Brian Ferguson

Working Paper No: 160006

www.canadiancentreforhealtheconomics.ca

May 19, 2016

Canadian Centre for Health Economics
Centre canadien en économie de la santé
155 College Street
Toronto, Ontario

CCHE/CCES Working Paper No. 160006
March, 2016

How does Insurance affect the Price of Drugs: A Graphical Analysis

Audrey Laporte[a,b,1], Brian Ferguson[b,c]

[a] *Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada*
[b] *Canadian Centre for Health Economics, Toronto, Canada*
[c] *Department of Economics and Finance, University of Guelph, Guelph, Canada*

Abstract

Prices of drugs differ greatly across countries and to a certain degree across payment agencies within countries (OECD (2015)). It is well known among health economists that the presence of insurance creates a separation between the consumer of pharmaceuticals and the payer. This separation can result in the price of drugs being driven up simply because somebody other than the consumer is responsible for paying for them. The precise impact of insurance on drug prices however, will depend critically on the structure of the insurance, a fact that has tended to get lost in health policy debate. The purpose of this paper is to use diagrammatic analysis of three types of insurance: co-insurance, reference pricing and co-payment, to investigate how each affects the price of prescription drugs. In addition, we analyze the role of a new pricing tool, which has recently been increasingly used by pharmaceutical companies in North America: co-Payment waiver coupons. Among other policy implications, we suggest that the use of co-pay waivers turns the co-payment insurance constraint into something similar to the reference pricing constraint, from the supplier's perspective, but with greater transactions costs.

Keywords: drug pricing; insurance; pharmaceuticals; co-payment; reference pricing; co-insurance

[1] Corresponding Author. audrey.laporte@mail.utoronto.ca

I Introduction

Prices of drugs differ greatly across countries and to a certain degree across payment agencies within countries (See, for example, OECD (2015)). The price of drugs is a contentious issue in policy debate but that debate frequently proceeds without a well-defined analytical framework. Many of the issues which seem particularly knotty in various countries policy debates can be understood using the techniques of microeconomic theory. For example, it is well known among health economists that the presence of insurance creates a separation between the consumer of pharmaceuticals and the payer. This separation can result in the price of drugs being driven up simply because somebody other than the consumer is responsible for paying for them. The precise impact of insurance on drug prices however, will depend critically on the structure of the insurance, a fact which has tended to get lost in health policy debate. A crucial factor is that the particular pricing strategy adopted by pharmaceutical companies will be determined ultimately by the nature of the insurance system within which they are working. Pharmaceutical companies are of course, profit-maximizers and will choose policies that maximize their profit subject to the institutional constraints under which they operate. Differences in the prices of drugs across countries can be related to differences in the insurance structure across those countries. Broadly speaking, health economics textbooks recognize three types of insurance: co-insurance, reference pricing and co-payment. These types co-exist to different degrees in different countries, and as we illustrate below turn out to have very different implications for the pricing strategies adopted by pharmaceutical companies.

It is generally understood that the structure of insurance can have quite significant implications for the behavior of various agents in the health care system, and hence for health spending[2], but the particular implications of each of the three types of insurance for the level drug prices is less well understood.

The purpose of this paper is to set out clearly the analytics of three types of insurance: co-insurance, reference pricing and co-payment, and consider how each affects the price of prescription drugs[3] In addition, we analyze the role of a new pricing tool, which has recently been increasingly used by pharmaceutical companies in North America: co-Payment waiver coupons.

While the properties of the three types of insurance can be set out formally in mathematical terms, their policy implications are best illustrated in diagrammatic terms. In each of the following sections we therefore use diagrammatic exposition to set out the essence of the argument, following which we shall compare certain key policy implications.

II Co-Insurance

Co-insurance refers to the case where the insured individual makes an out-of-pocket payment which is calculated as a percentage of the market price of the health care product being purchased. This percentage is referred to as the co-insurance rate. thus in the case of a 10% co-insurance rate, the patient pays 10% of the price out of pocket: if the drug has a market price, set by the producer, of \$100, the patient pays \$10 out of pocket and the insurance plan pays \$90. If the market price rises

[2]See, for example, Comanor and Schweitzer (2007)

[3]In our analysis we treat the supply side as completely vertically integrated and do not deal with pricing between the manufacturing, wholesale and retail sectors.

to $200, the patient's share rises to $20 and the insurer's to $180. The idea behind co-insurance is that the explicit tying of the patient's out of pocket payment to the level of the retail price should make the patient sensitive to the price of the drug, even if he is not paying the full thing. This in turn was presumed to impose demand side restraint on the price-setting decisions of the suppliers. Note that in this form of insurance the insurer does not set the price that it will pay for the product - the usual pattern is for the insurer not to interfere in the price setting process.

We can see the formal effect of co-insurance if we start with the case in which there is no insurance, and introduce insurance of the co-insurance form. In the pre-insurance case, if we assume, for simplicity, a monopoly supplier of the drug, pricing will obey the usual monopoly pricing rules. This means that the supplier will face an Average Revenue (AR) curve which is identical to the consumers' Willingness to Pay (WTP) curve at the level of the market. The market WTP curve is simply an aggregation of the WTP values of the individual consumers - these in turn reflect the value of the marginal benefit (VMB) which each consumer expects to derive from the product. Normally we would refer to this aggregated VMB curve as the market demand curve. Part of our purpose here, however, is to illustrate that our understanding of certain policy conundrums will be greatly enhanced if we change the perspective from which we look at the market in question. In the present case, it helps to look at the demand side of the market explicitly in terms of prices received by producers (average revenue) and the prices which consumers judge to be the highest they are willing to pay for successive units of the commodity (WTP or Demand Prices).

The pre-insurance monopoly market is set out in Figure 1 below:

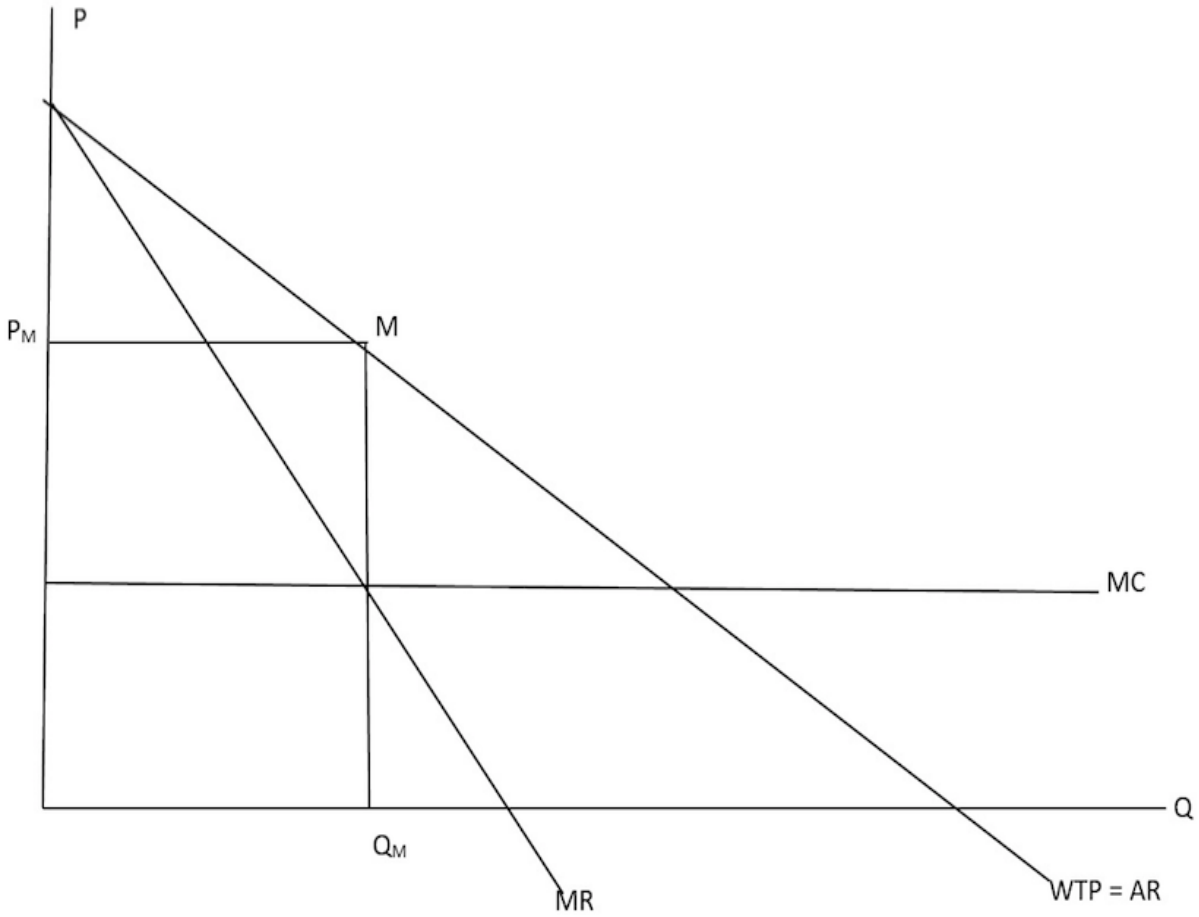

Figure 1: Pre-Insurance Monopoly

Figure 1 is the standard textbook monopoly diagram, with the monopoly price and quantity determined by the intersection of the MR and MC curves. In this diagram, since we are assuming that the supplier is a single price monopolist, the WTP curve is also the firm's AR curve.

Next, consider the introduction of a co-insurance rate c. With price P_M, the amount paid out of pocket by the consumer will be cP_M, and the amount paid by the insurer will be $[1\text{-}c]P_M$. This gives Figure 2 below:

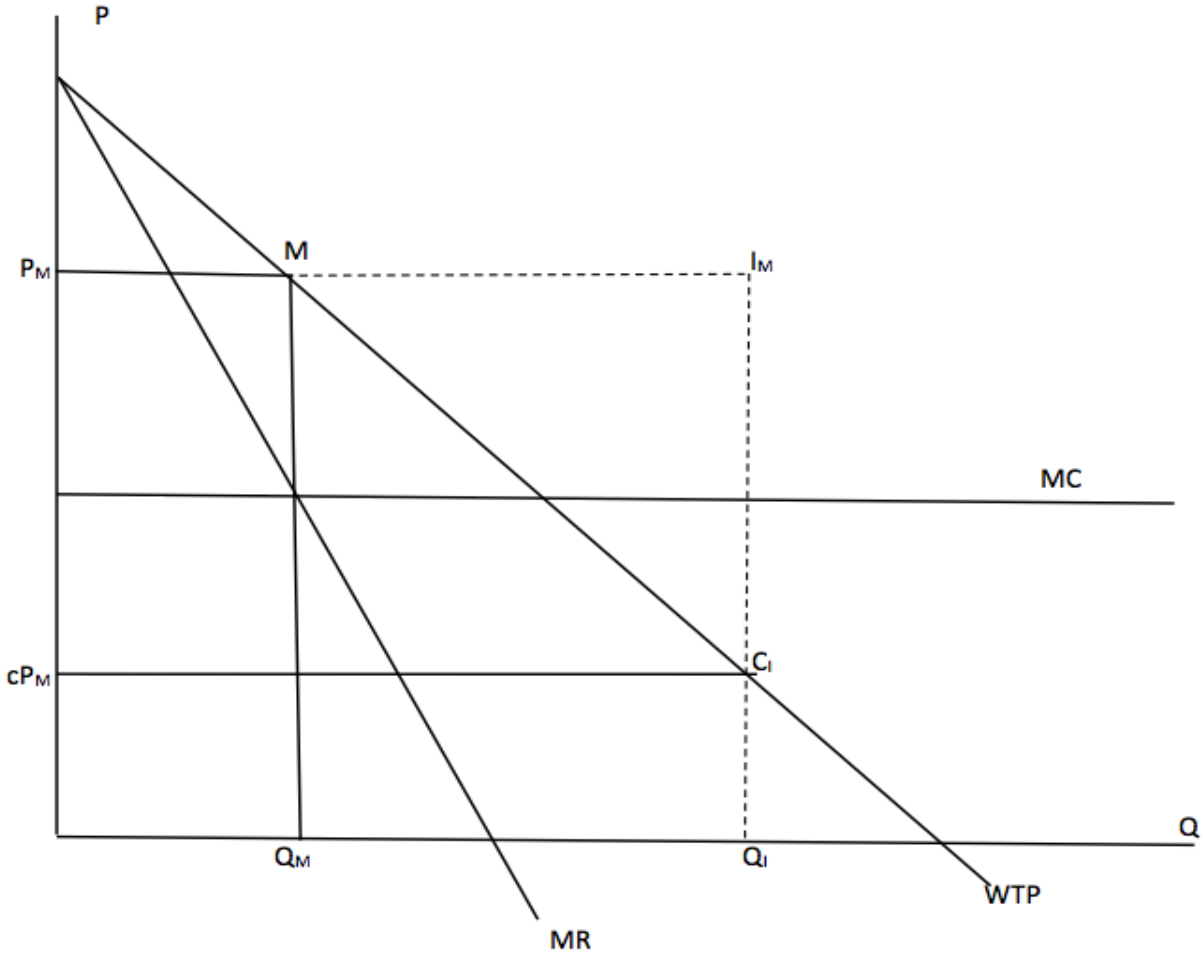

Figure 2: Introduction of co-insurance

Figure 2 is the usual co-insurance diagram. The price received by the supplier remains unchanged at P_M but because the price paid out of pocket by the patient falls to cP_M, the quantity demanded rises to Q_I. The consumer's optimum is now at point C_I. Because the consumer's marginal benefit at C_I, equal to cP_M, is below the (assumed constant) marginal cost of quantity Q_I, this diagram is typically used to illustrate the welfare loss from excess health insurance[4]. It is important, however, that we not stop at this point. The fact that the presence of insurance puts a sort of wedge between the demand and supply sides of the market means that it is important that we look at the market for pharmaceuticals in the presence of insurance both from the consumer's perspective and from the producer's perspective. It is, after all, the supplier who will be setting the price[5].

[4]The classic reference on excess health insurance is Martin Feldstein (1973).

[5]Suppliers in imperfectly competitive markets have pricing power, but it is important to remember that that power is always constrained by the consumer's willingness to pay for different quantities of the good.

If we look at this market from the perspective of the supplier, we see that, while the price it receives per unit (its Average Revenue per unit) has remained unchanged at P_M, the quantity which it is selling at that price has increased from Q_M to Q_I. Since this price-quantity pair must be (by definition) on its AR curve, and since the horizontal intercept of its AR curve does not change (that being the quantity which is so large that the next unit yields a marginal benefit of zero to the consumer) the supplier can trace out a new AR curve through the horizontal intercept of the consumer's WTP curve and this observed point on the supplier's new AR curve. This is shown in Figure 3 below[6].

[6]For this diagram see also Berndt, McGuire and Newhouse (2011).

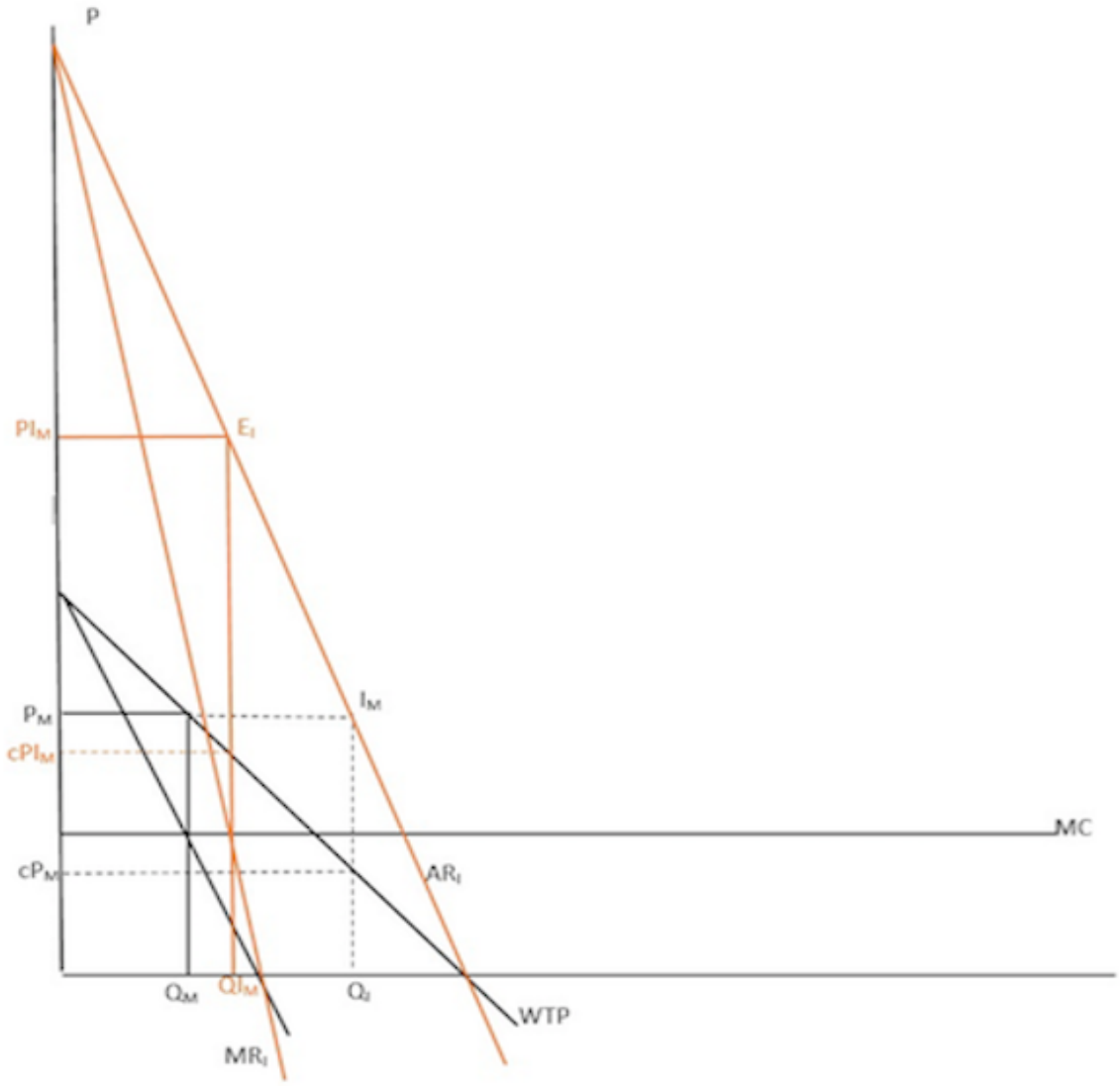

Figure 3: Effect of co-insurance on drug retail price

The initial effect of the introduction of co-insurance is to increase the quantity demanded from Q_M to Q_I. That is not, however, the end of the story. Looked at from the supplier's point of view, the AR curve has rotated out, to go between the horizontal intercept of the patients' WTP curve and the point determined by P_M and Q_I. The rotation of the AR curve, however, also implies a rotation of the supplier's MR curve, again when we look at the market from the supplier's perspective. This in turn means that the original profit maximizing point is no longer a profit maximizing price/quantity combination, and this in turn causes the supplier to begin to raise the price.

The process is never expressed in quite those terms - traditionally the outward rotation of the AR curve has been described as "patients no longer being so constrained financially in terms of the amount of care they take". More recently the process has been characterized more explicitly, with statements to the effect that insurance cushions patients from price increases. All of that, however, is public relations - in practical terms what has happened is that the supplier has realized that it can sell more units at the old price, which suggests that raising its price, while it will cost it some sales, will not be as damaging as would have been the case before the co-insurance system cut the out of pocket price to the patient. As the price drifts up, in fact, the supplier will find its profits increasing. It will continue to raise the price until it reaches the point on the new AR curve which is associated with intersection between the new AR curve and the associated new MR curve. This point will be at price PI_M and quantity QI_M.

We note that while PI_M is much higher than the original price, P_M, QI_M is only slightly higher than the original quantity point, Q_M. This is important to our understanding of the dynamic involved. The quantity of care actually consumed is determined by the patient, in a process which is summarized in the WTP curve. The fact that the new long run profit maximizing quantity is only slightly larger than the original quantity means that the amount which the patient is being asked to pay out of pocket per unit of quantity must be only slightly lower than the original, pre-insurance out of pocket price to the patient. On Figure 3, the new long run out of pocket price is cPI_M which is only slightly below the original price P_M.

The long run effect of co-insurance, then, will be to raise the market price of the drug to well above its non-insurance monopoly level but to increase utilization - usually referred to as access to care - only slightly, and, in terms of payment out of pocket, to leave consumers in virtually the same position as they were in before the introduction of insurance. Co-insurance, which was intended to increase access and at the same time restrain prices by ensuring that consumers continue to have skin in the game, effectively serves as a mechanism to raise the full market price of the drug while, in the long run, leaving access to care virtually unchanged[7].

Our analysis of the effects of co-insurance rests on the recognition that we must take account of the two perspectives which are at play in this form of market. The first is the consumer's perspective of the benefit he will derive from consuming the drug (or other treatment) in question, which is reflected in the WTP curve. The second is the supplier's perspective on the price that it will be able to charge per unit for different quantities of its product, as summarized by its AR curve. Normally the two curves coincide, but one of the key features of the market for health care is that the AR curve is separated from the WTP curve in a manner, and to a degree, which depends on the type of insurance involved.

Note that we have said that there are two perspectives involved in the analysis - that of the patients and that of the suppliers. We have not considered the perspective of the insurer. That is essentially because, apart from determining the co-insurance rate, the insurer has tended to play a relatively passive role in co-insurance systems. In recent years insurers have played a more active

[7]The case of Claritin provides a useful illustration of this effect. There was a period in the 1990s when Claritin was a prescription drug in the United States and an Over the Counter Drug in Canada. As a prescription drug it was covered by insurance, as an OTC it was not, and accordingly its retail price was much higher in the US than in Canada - in terms of Figure 3 the retail price in the US was at PIM while in Canada it was at PM. When Claritin lost its prescription status in the US, the price of a box of Claritin fell but there were complaints that the price to American consumers had risen. This was, in terms of our Figure, an increase from CPIM to PM, resulting from the collapsing of the higher of the two AR curves into the lower one. See Freudenheim, (2003).

role in co-payment and reference pricing systems. We next consider the formal analytics of these systems. In both of them the need to consider the AR and WTP curves as representing separate perspectives continues: the difference between these systems and the co-insurance system will come down to the type of gaps which these different insurance structures insert between the AR and WTP curves.

III Reference Pricing

Reference pricing as an insurance system is probably more widely known in European health insurance systems than it is in North America, although it has been used in British Columbia in the public drug insurance program, and in California in certain hospital insurance programs[8]. It is also used in some employer-sponsored health insurance plans, although generally not under that name - many employer-based plans for example, may specify an amount which they will pay for eyeglasses every two years but leave the actual pricing of the glasses to the judgment of the optician. This will, of course, be a market judgment by a profit-maximizing supplier.

Under what we might term a generic reference pricing system, the insurer defines a category of product - we shall continue to use prescription drugs as our example - and sets the price which it will pay for any member of the category. It is important to note that the insurer is paying exactly the same price for each member of the pricing group, it is not setting a different price for each. The actual market price of each member of the group is set by the supplier, and the difference between the reference price, which is paid by the insurer, and the market price is paid out of pocket by the patient. Thus the patient's WTP will come into the determination of the profit-maximizing price for the supplier.

At one level the formal analysis of a reference pricing system is easier than that of a co-insurance system. The supplier's AR curve still differs from the patient's WTP curve, but the difference is a constant amount per unit, equal to the reference price (which is defined per unit) because the reference price does not differ with the number of units bought or with the price per unit charged on the market by the supplier. Thus we can represent a reference pricing system as a combination of two graphs, one the patient's WTP curve and the other a constant reference price (RP) line. In Figure 4 below shows the two elements individually:

[8]See, for example, Robinson and Brown (2013). See also Pear (2016) for a newspaper report referring to consideration of using reference pricing in US Medicare.

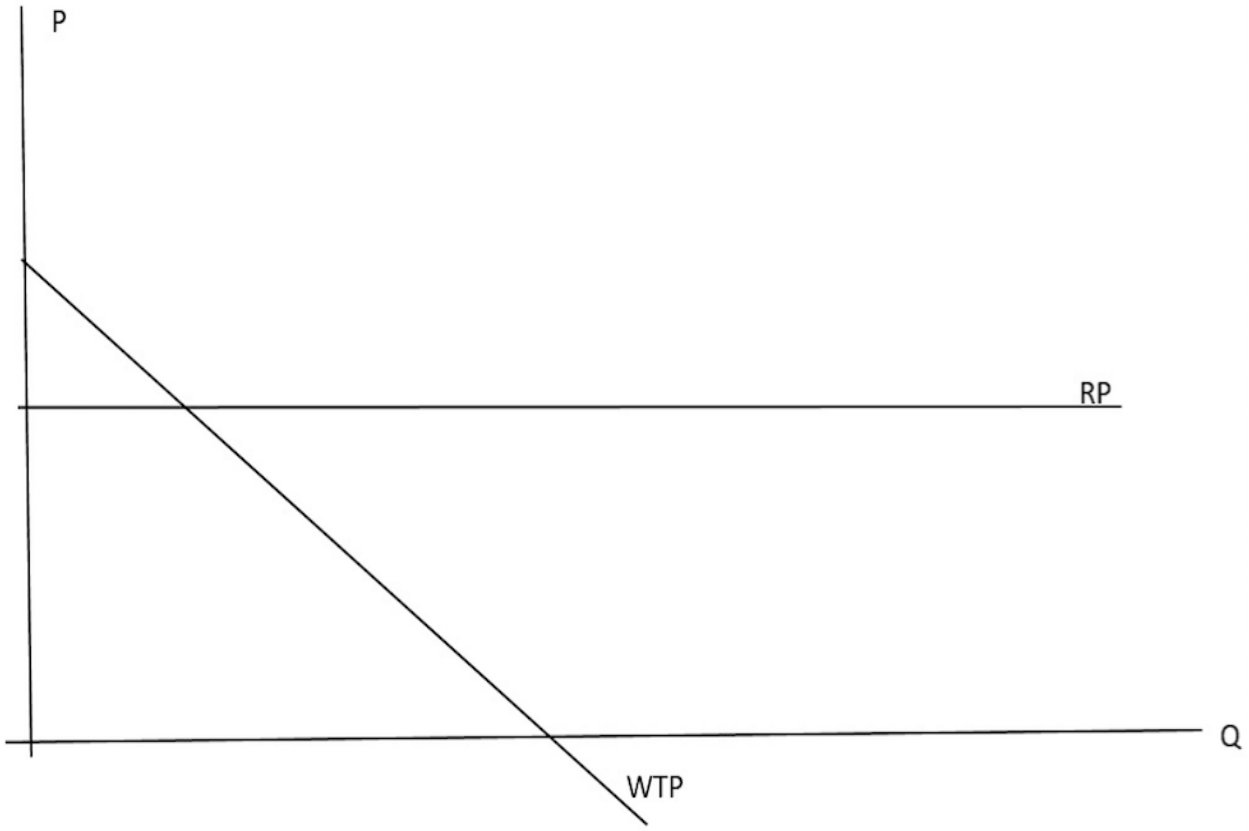

Figure 4: Elements of the Reference Pricing system

Formally, the AR curve will be a vertical summation of the two elements shown in Figure 4. This is because the firm's total revenue will be equal to [RP + P(Q)]Q, where RP is the reference price and P(Q) is the patients WTP per unit for Q units, the fact that there is functional notation on the WTP element but not on the RP element reflecting the fact that the WTP amount varies with the quantity which the patient is considering buying while the RP amount does not. AR at quantity Q, then, is just [RP + P(Q)] and marginal revenue is [RP + P(Q)] + P'(Q)Q, where P'(Q) is the slope of the WTP curve. In terms of the diagram, the easiest representation involves deriving the AR curve by seating the WTP curve on top of the RP line, not losing sight of the fact that this represents the supplier's perspective: to see the patient's perspective in this diagram we need to treat the RP line as the horizontal axis for the WTP diagram as in Figure 5.

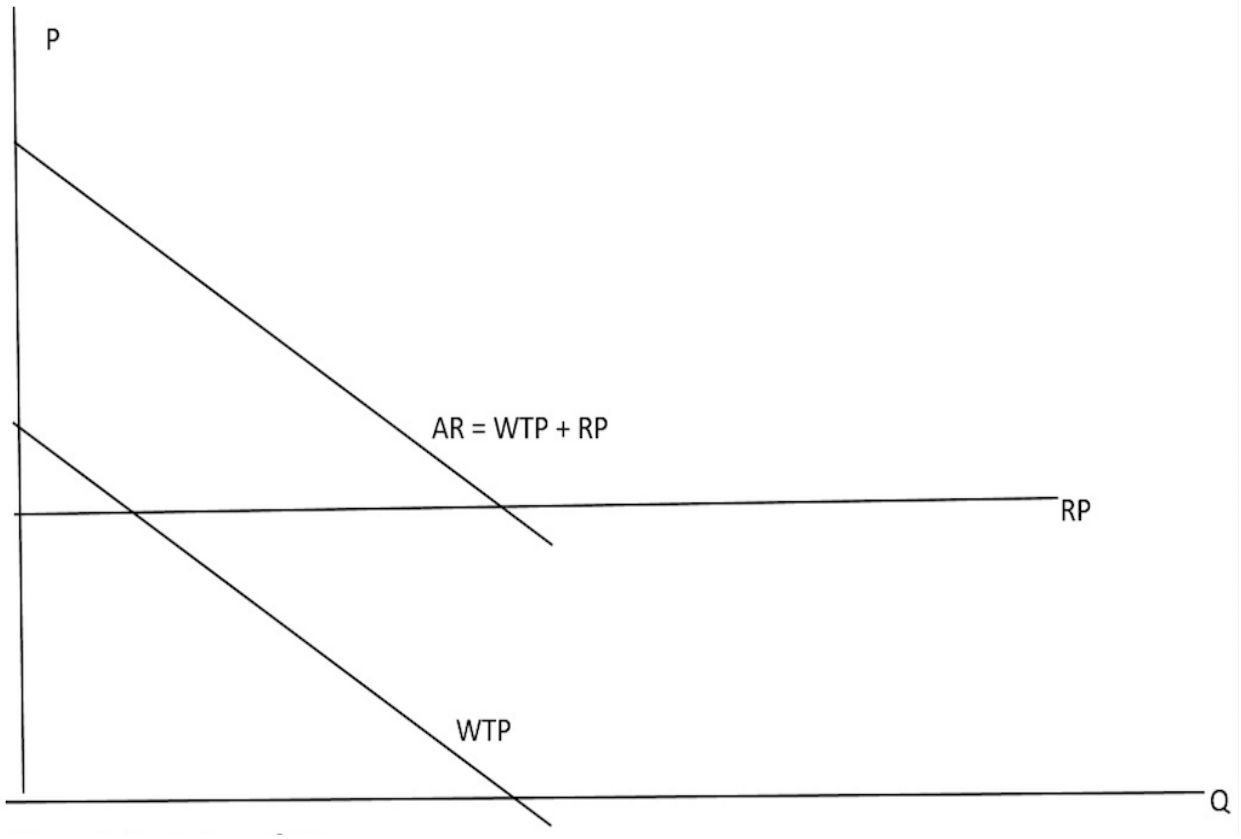

Figure 5: Derivation of AR curve

Along with the AR curve there will be a MR curve, shown in Figure 6 below where we have also added the Marginal Cost (MC) curve:

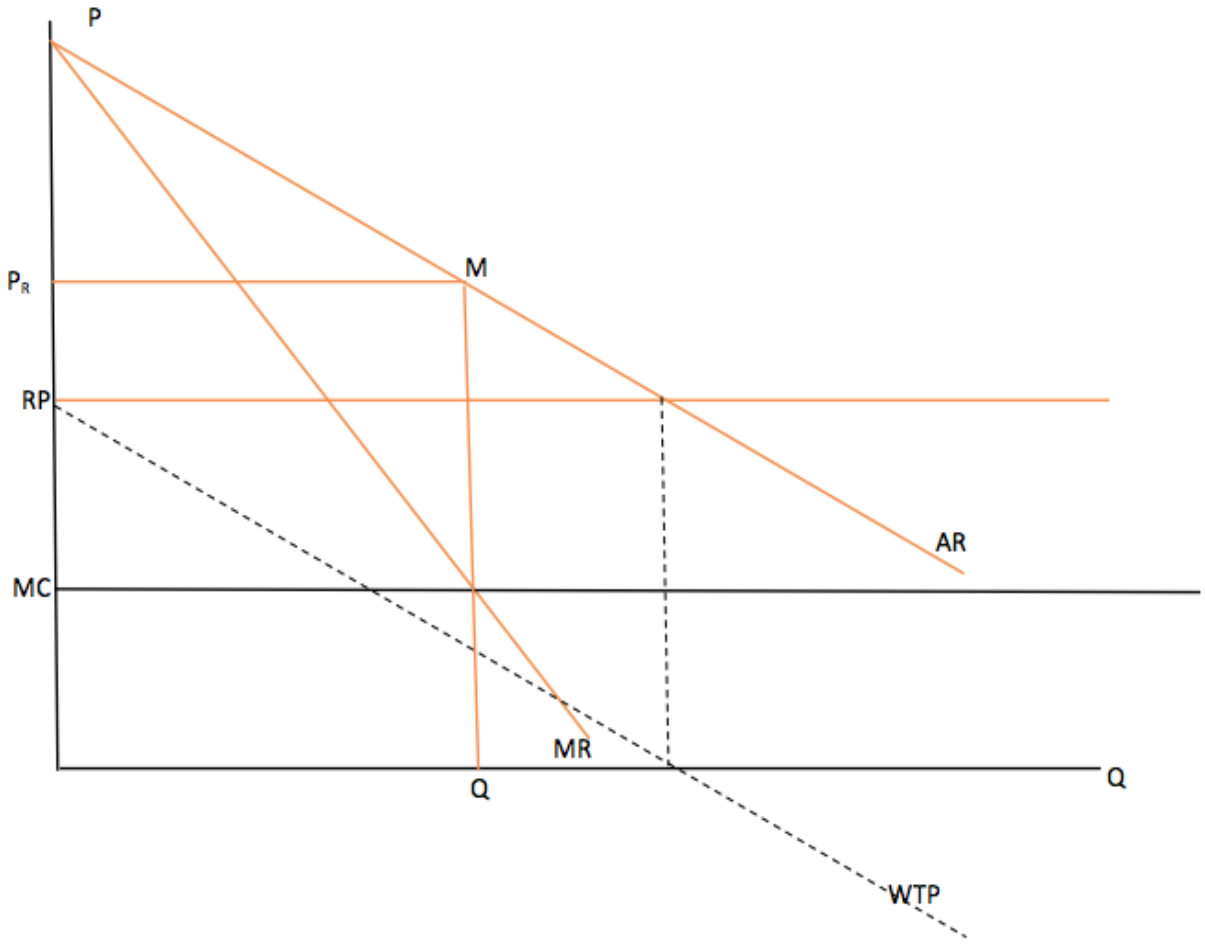

Figure 6: Reference Pricing

In Figure 6 the full (profit maximizing) market price of the drug is P_R, determined at the intersection of the firm's MR and MC curves. We note that, while the firm collects the price PR per unit, RP of that is paid by the insurer and the consumer pays P_R - RP out of pocket. This is why we said above that, in this diagram, to see the patient's perspective once we have shifted the WTP line up as part of the process of deriving the AR line, we need to treat the RP line as the horizontal axis for the WTP diagram. (We could work with the WTP curve positioned relative to the true horizontal axis, as shown by the dashed WTP line in Figure 6, but the diagram as set out with the solid lines in Figure 6 is more useful for the policy discussion in the later sections. Note that the dashed WTP line is parallel to the AR line, since the AR line is the dashed WTP line shifted up by a constant amount RP at each value of Q.)

We have drawn Figure 6 from the perspective of a single firm which supplies one of the items which

has been included in this particular reference pricing group. This means that the WTP curve which we have drawn represents the consumer's WTP out of pocket for that particular firm's product. This in turn will depend on the number of close substitutes to be found in the reference group - the more close substitutes there are the more elastic will be the WTP curve for any individual supplier's product, even if the class which defines the group would be classified as a necessary medication, in the sense that the demand for the active ingredient in the medication would be highly price inelastic. The essence of the Reference Pricing form of insurance lies in the difference between the elasticity of demand for individual elements in the reference group and the elasticity of demand for the active ingredient which defines the group itself. The more individual products there are in a single group and the closer substitutes they are from the perspective of the patient, the flatter will be the WTP curve for any individual element in that group and the lower will be the patient's out of pocket payment. This last result, about the out of pocket price, follows from the fact that, because the AR curve is a vertical shift of the dashed WTP curve and is therefore parallel to the WTP curve, the AR curve will cut the RP line directly above the point where the dashed WTP curve cuts the RP curve. As the WTP curve becomes more elastic it will rotate on that point, pivoting so that it becomes flatter, but not tend to shift. Thus the horizontal intercept of the WTRP curve and the horizontal coordinate of the point at which the AR curve cuts the RP line become fixed points for the analysis.

We should note that Reference Pricing allows there to be differences in the market price, and hence the out of pocket price, across drugs, on the basis of differences in the elasticities of demand for the individual elements within the group. In particular, if a significant number of patients have a strong preference for the brand name drug, its individual demand curve will be less elastic than that of its generic competitors and it will be able to set a higher price than they could; the difference in effect reflecting differences in patients' strengths of preference.

Our discussion of Reference Pricing has assumed that there are, in the class under consideration, a number of competing products. This could include several brand name drugs, which have a similar mechanism of action (although their mechanisms of action must differ sufficiently that they do not violate each other's patents) or, more simply, the class could consist of an off-patent brand name drug and its generic competitors. There are obvious policy issues here relative to the setting of the Reference Price, but we will set them aside for discussion in the policy section of this paper. Before proceeding to that, we shall discuss a formal graphical representation of the co-payment form of insurance.

IV Co-Payment

The basic co-payment model has a lot in common with the RP model, but is sufficiently different that its effect on the pharmaceutical market will be quite distinctive. As in the RP case, the retail price of a drug is set by the supplier, and the patient is required to pay a certain amount out of pocket. Under co-payment, however, the price set by the supplier is the upper limit of the payment the supplier receives for their product. When the patient fills a prescription for a covered drug, the insurer pays all but the co-pay, which is collected directly from the patient. The amount of the co-pay is set by the insurer, with an eye to influencing the patient's use of the drug or its substitutes, if any.

In terms of a diagram, we again show a horizontal line representing the price of the drug, and a negatively sloped line representing the patient's WTP curve, but this time, as in Figure 7 below, we position the WTP curve on the horizontal axis, not on the F-line.

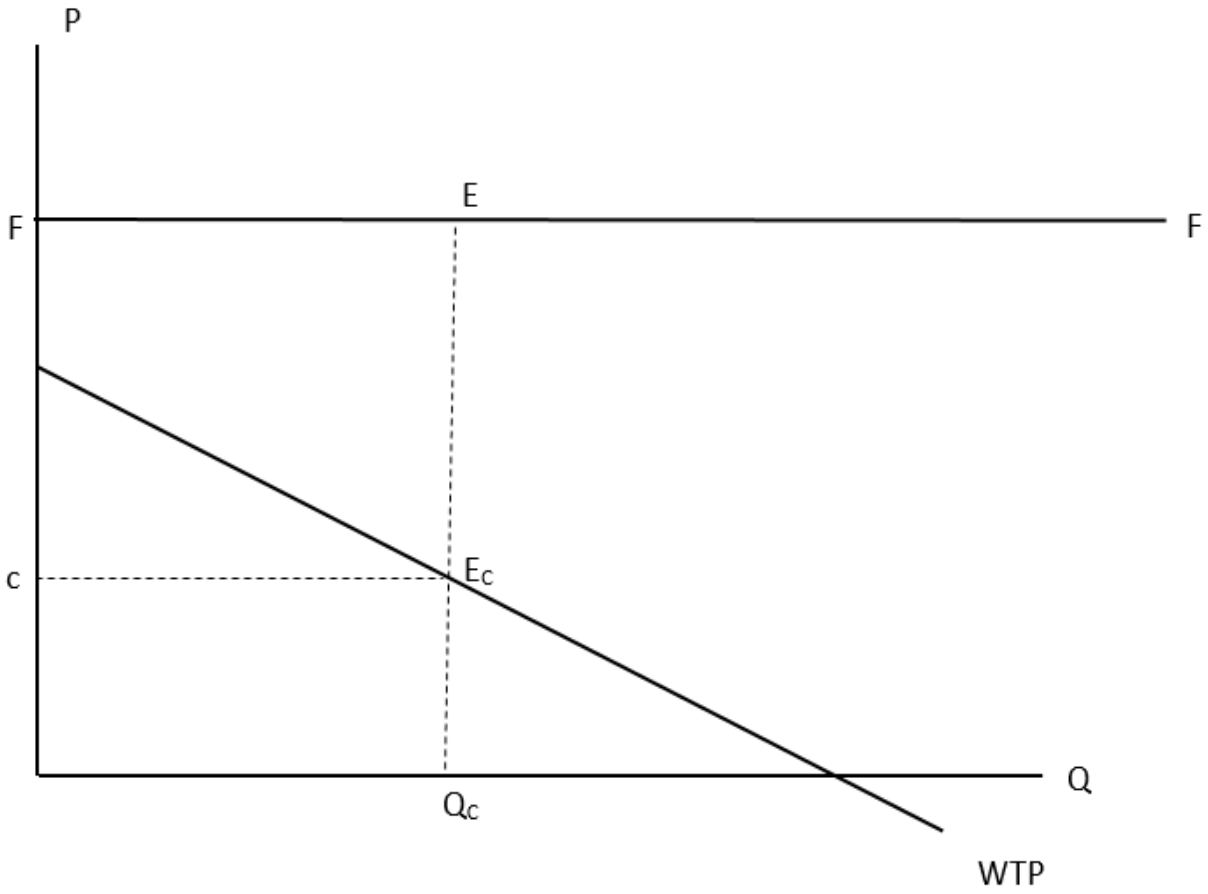

Figure 7: Co-Payment

In Figure 7, while F is the full price of the drug, c is the co-pay amount which the patient is expected to pay out of pocket and Q_c is the quantity of the drug that will be demanded at a co-pay of F. The amount F-c is the per-unit price which will be paid by the insurer. The total expenditure on the drug will be the area of the rectangle defined by F, E Q_C and the origin, while the amount paid out of pocket by the patient will be the rectangle defined by c, E_C, Q_C and the origin, with area F, c, E_C,E paid by the insurer. Thus in the co-pay model the insurer makes use of the patient's WTP curve to influence the quantity of the drug demanded and the amount the insurer has to pay.

Co-pay plans typically make use of co-payment tiers to try and shift patients between substitute drugs. In considering substitutability, insurers are effectively assigning drugs to therapeutic categories, as in the case of reference pricing. This means that when the same drug could be used to

treat several different conditions, an insurer could make the co-pay for that drug diagnosis specific, depending on the set of substitutes available for treating each diagnosis.

We consider here the case where there is one brand name drug and one generic competitor, and where the generic competitor has set a lower price for their drug than has the brand name supplier[9].

[9]Previously, the literature has assumed that the supply curve of generics is horizontal and that the market for generics will ultimately be perfectly competitive, even if there is a single first-entrant generic producer for a few months. This structural assumption was implicit in the assumption that the price of generic drugs would automatically be significantly lower than the price of brand name drugs, being roughly equal to the constant unit cost of the cost of the physical production of the pills. Recent American experience has indicated that this assumption was too sanguine: if a generic company can acquire a monopoly, even for a short period, it can be expected to exploit it for as long as it takes for competitors to enter. To the extent that regulatory requirements can impose barriers to entry into a generics market, there may be an opportunity for monopoly pricing for several years. To a certain degree this situation is a consequence of the constant returns nature of the physical production process - at one time the rate of profit in the generics sector was greater than that in the brand name sector. That profit attracted competitors in, driving the price of generics down to the level of the marginal cost of production. It seems likely that the response to this situation was for suppliers to exit certain markets, creating monopoly or duopoly supply situations, which in turn led to significant increases in drug prices. The resultant high profits could not lead to quick entry, partly because of the fixed costs of establishing production facilities but also, in the US, because of the regulatory approval lags associated with the entry of a new generics product and production facility.

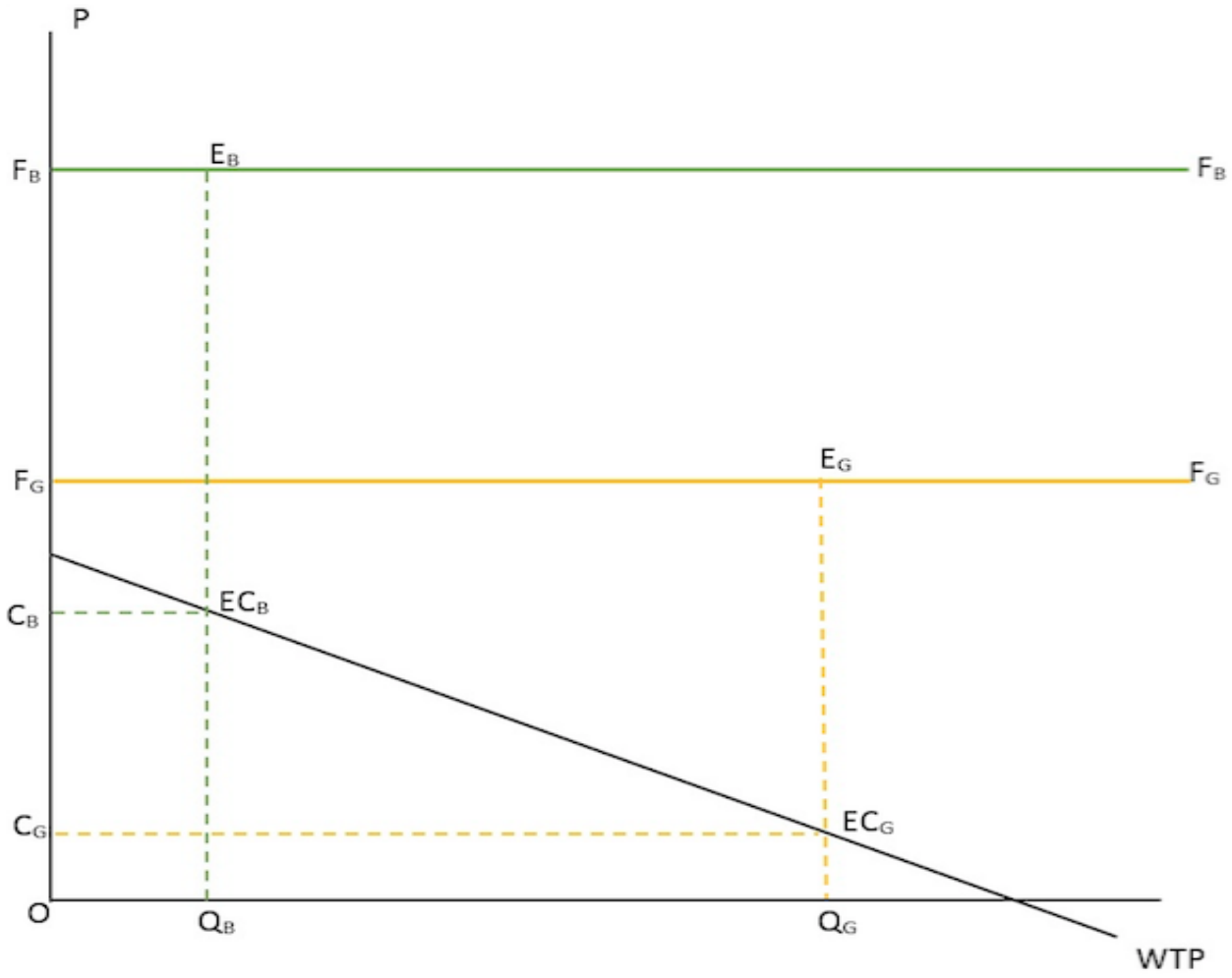

Figure 8: Brand name and generic drugs

In Figure 8, for convenience, we assume the same WTP for both drugs, reflecting the assumption that the generic is a very close substitute for the brand name. Here the patients who choose the brand name drug, B, must pay C_B out of pocket and those who choose the generic pay C_G. The difference between the co-pay amounts is aimed at being large enough to ensure that the only patients who choose the brand name drug will be ones who feel very strongly about staying with it; the rest will shift to the generic. In Figure 8, the quantity taken by patients who stay with the brand name drug will be Q_B, and the insurer's payment for the brand name drug will be the area of the rectangle F_B C_B EC_B E_B, while the total paid by patients for B will be the area C_B EC_B, Q_B,O, based on an out of pocket co-pay of C_B per unit. In the case of the generic drug, G, the quantity demanded by patients will be Q_G, the amount paid by the insurer will be F_G, E_G, EC_G, C_G, and the total amount paid by patients will be C_G EC_G, Q_G, O, based on an out of pocket co-pay of C_G per unit.

As we noted above, the purpose of co-pay tiers is to allow the insurer to shift patients from more expensive brand name drugs to less expensive generics, when the generics become available. Typically when there is only one drug, a first entrant brand name drug, in the market, the drug will be put in a lower co-pay tier, and shifted to a higher tier when generics come on stream (there may also be a shift in tier if a competing brand name comes on stream and is priced significantly below the first entrant drug.) Thus the insurer is making use of the elasticity of the consumer's WTP curve to shift consumption between drugs.

V Co-Pay Coupons

The co-pay diagram can be used to analyze a policy recently adopted by a number of brand name drug companies after generic competitors for the brand name drugs have entered the market, of offering co-pay coupons to patients who stay with the brand name drug[10].

In Figure 9 below we reproduce Figure 8 as the brand name firm sees it, i.e. removing the generic drug from the diagram. As Figure 9 stands, the Brand name supplier will receive total revenue equal to F_B, E_B, Q_B, O, receiving the upper rectangle from the insurer and the lower one from the consumer. Co-pay coupons are designed to affect the price paid by the consumer without altering the price paid by the insurer. Thus the aim is to keep F_B C_B unchanged while changing the quantity of the brand name drug which is demanded by altering the amount which the patient pays out of pocket.

This is accomplished by issuing coupons to patients whose doctors write prescriptions for the brand name drug, and who forbid substitution of a generic for the brand name. The coupon represents a discount of the co-pay which the consumer would normally pay out of pocket[11].

[10] See Ross and Kesselheim (2013) for a discussion of the rapid recent growth in the use of co-pay coupons.

[11] In our discussion, the patient is assumed to be buying the drug directly from the brand name company. Clearly this is an oversimplification: a full analysis would need to allow for the role of the pharmacy at which the patient has his prescription filled. Roughly speaking, the pharmacy buys the drug from the supplier - either the manufacturer or a wholesaler - in the way that retail operations in general acquire the products which they sell to customers. In some cases the patient would pay the full amount FB out-of-pocket, submit a bill to his insurer, and be reimbursed the insurer's portion of the cost, leaving the patient out-of-pocket for the co-pay. In others the insurer pays its portion of the cost directly to the pharmacy, leaving the pharmacy to collect the co-pay amount directly from the patient. When a co-pay coupon has been issued by the brand name firm, the pharmaceutical manufacturer pays, all or part of the co-pay to the pharmacy on behalf of the patient. If retail pharmacies were going to be out-of-pocket as a result of accepting co-pay coupons, we would expect them to decline the coupons.

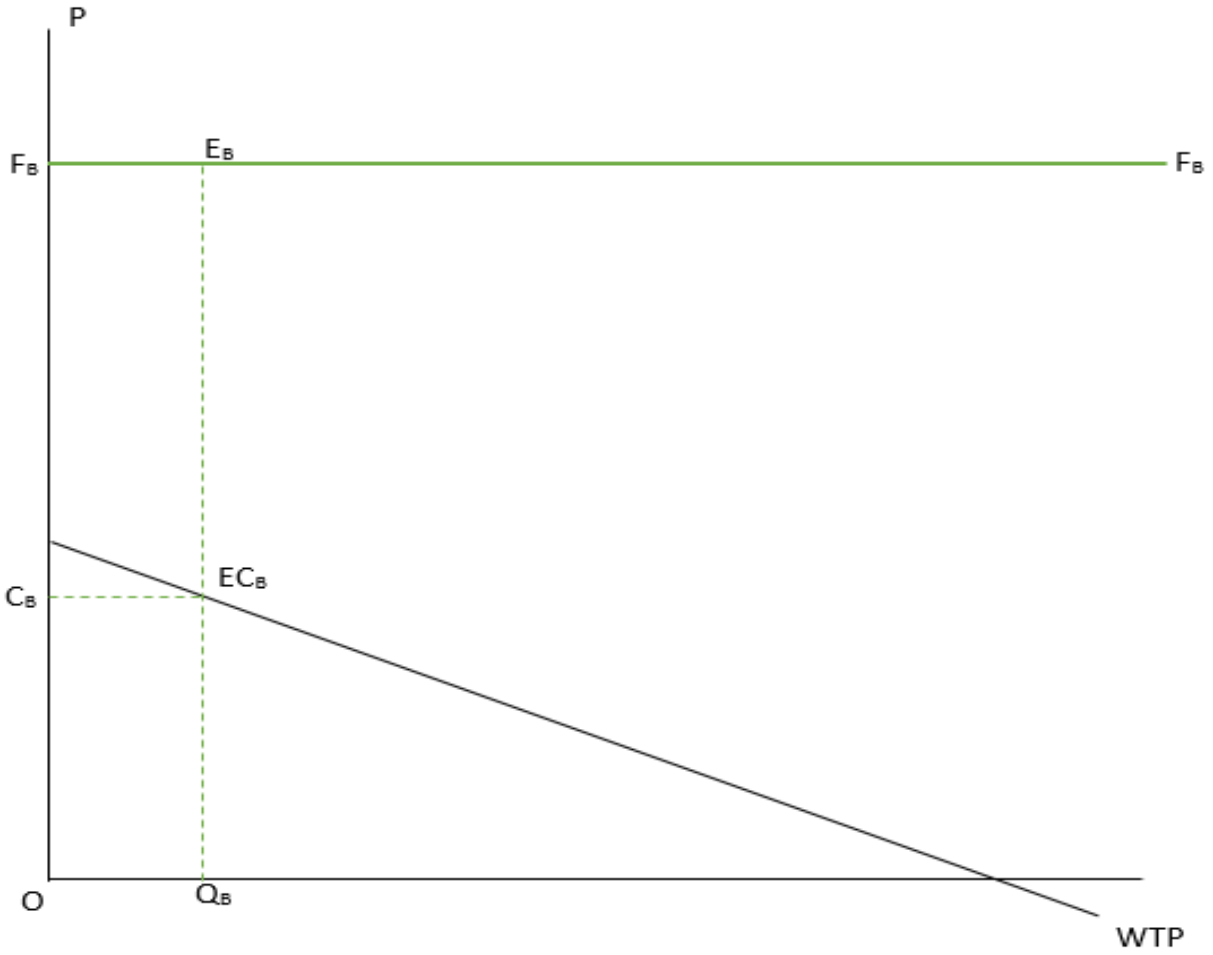

Figure 9: Brand Name Co-Pay

By issuing the coupon the manufacturer is effectively waiving part of the out-of-pocket payment, reducing C_B in the eyes of the consumer. The insurer has determined the amount it will pay on the assumption that the co-pay will be C_B, and will not increase its amount simply because the patient is paying less - thus the insurer will continue to pay F_B - C_B per unit of quantity, even when the co-pay has effectively been reduced to C_B - R where R is the "rebate" granted to the consumer by the manufacturer.

To see the effect of this in the diagram, note that however the arrangement is structured - whether the manufacturer is waiving some of the co-pay directly, or asking the pharmacist to waive it and paying the waived amount directly to the pharmacist (in effect re-paying the pharmacist part of the price per pill which the pharmacist paid to acquire her inventory) we can think of the effect as being to reduce the total price which the drug company receives per pill by the amount of the

value of the coupon. In terms of Figure 9 above, this in effect makes the F_B line downward sloping starting from E_B as in Figure 10 below:

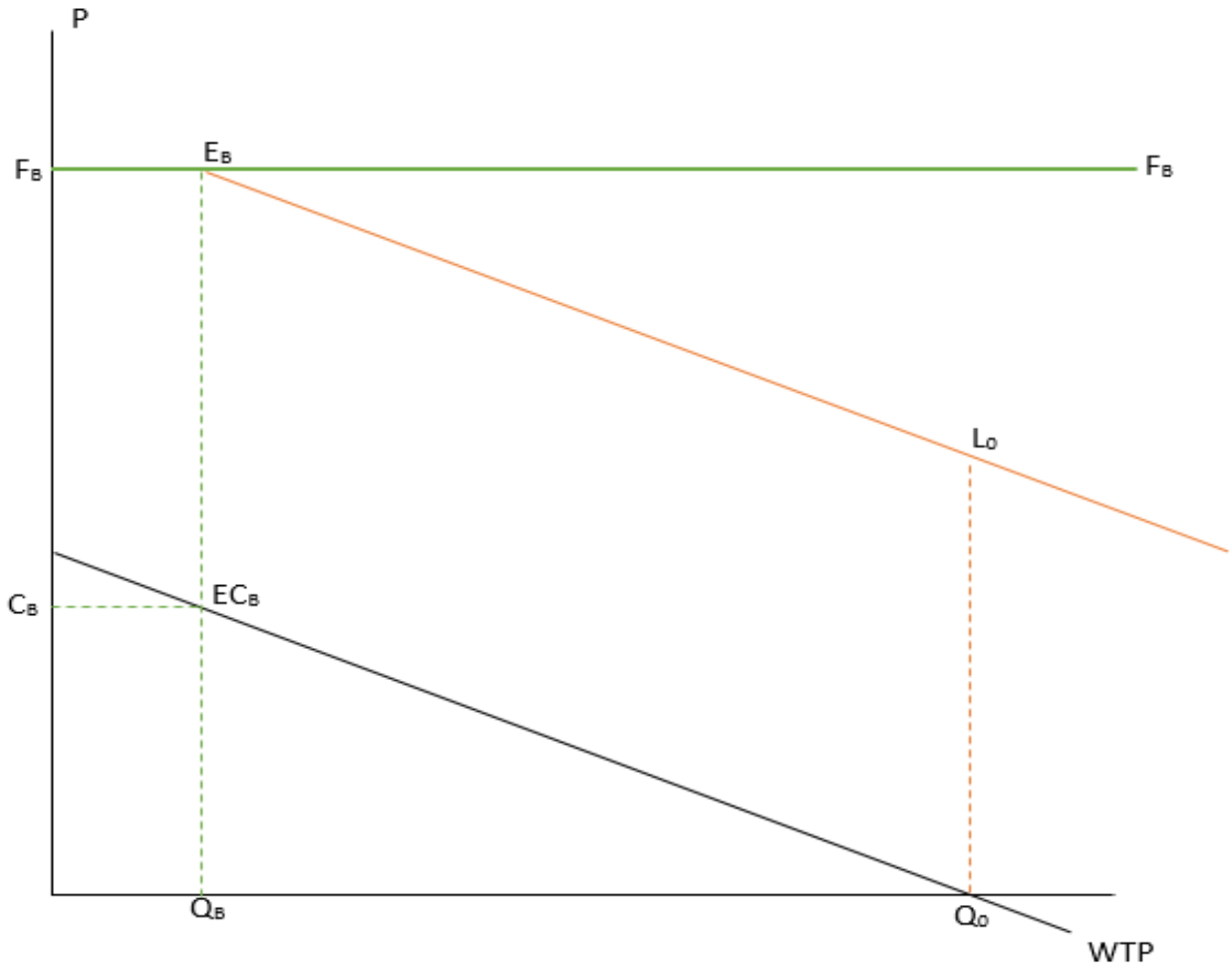

Figure 10: Co-Pay coupons

In Figure 10 the line F_B E_B L_0 is the brand name company's Average Revenue (AR) curve with the coupons. Q_0 is the quantity which patients would demand if their out of pocket price were zero, so the vertical distance between Q_0 and L_0 is the amount of the insurer's payment per unit of output. The new AR curve is identical to the F_B line up to EB, since the co-pay as set by the insurer is CB, then parallel to the original WTP curve along the segment E_B L_0.

Copayment insurance is intended to give the insurer control over costs by allowing it to adjust the price to the patient in order to influence quantity demanded. Clearly this requires some notion of the price elasticity of demand, although perhaps not an excessive amount - the use of tiers is

a fairly blunt pricing instrument. Tiered co-payment could also be seen as giving the insurer the opportunity to make consumers more sensitive to substitution possibilities than a constant-rate co-insurance system would, since the tiers can be used to make the out of pocket price rise faster than does the retail price, whereas in co-insurance the out of pocket price will in general be a constant percentage of the retail price and so rise in proportion to the retail price.

The recent introduction by producers of co-pay coupons, however, changes the nature of the tool. To understand how a profit maximizing pharmaceutical company can make use of co-pay coupons, we can add Marginal Revenue and Marginal cost curves to Figure 10, as in Figure 11 below. Since the AR curve is kinked, with a horizontal segment, the MR curve will be stepped: it will coincide with the horizontal segment from F_B on the vertical axis to point E_1, then drop vertically to point B, after which it will be downward sloped, twice as steep as the downward sloping AR_C curve.

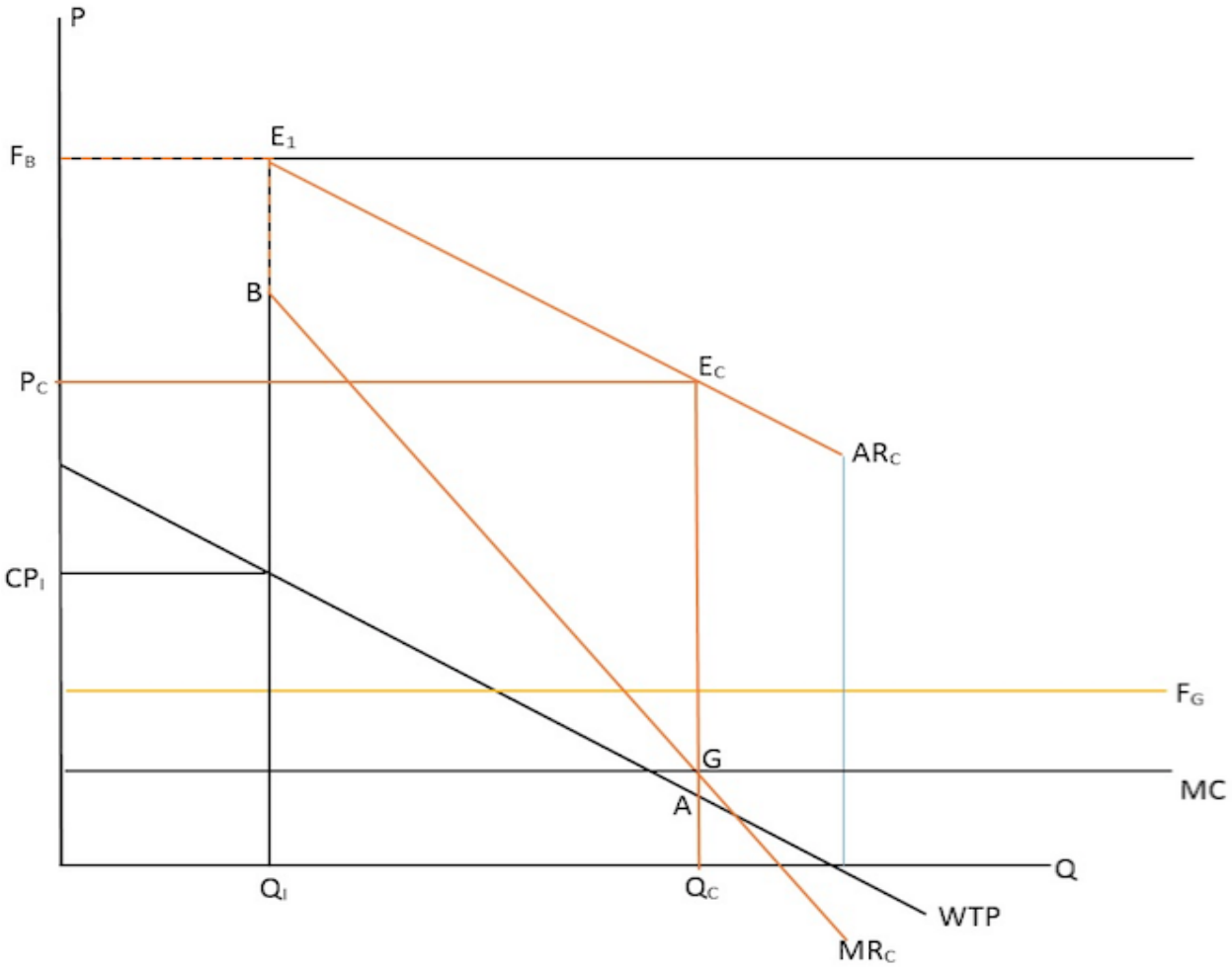

Figure 11: Profit maximization with Co-Pay coupons

The profit maximizing MR/MC intersection will be at point G, giving profit maximizing quantity Q_C. The profit maximizing point on the AR_C curve will be at E_C and the profit maximizing post-coupon price will be P_C. P_C will be less than F_B by the amount of the co-pay which has been waived. The vertical distance from E_C to G equals the distance from F_B down to CP_I on the vertical axis - i.e. the amount of the full retail price which the insurer had agreed to pay. This distance is also equal to the vertical distance from the horizontal intercept of the WTP curve to the AR_C curve - if the pharmaceutical company were to waive the co-pay completely, the quantity demanded would be at the horizontal intercept of the WTP curve and the total amount the firm would collect would be the point vertically above it on ARC, all of which would come from the insurance company. Note that, as we have drawn Figure 11, this vertical distance is greater than F_G, the full retail price of a generic drug, so the brand name company is better off not simply cutting its price to match the generic. The amount paid directly to the drug company by the patient (i.e. the non-waived portion of the co-pay) is the vertical distance from Q_C up to A, the point on the WTP curve directly above the profit maximizing quantity.

Suppliers can in effect, treat the insurer's portion of the payment as fixed, as in a two-part pricing scheme, albeit one in which the fixed element is not entirely under the supplier's control. We say "not entirely" because the supplier actually has two instruments open to it: in addition to the co-pay coupon, the price F_B which it sets as the retail price will determine the tier to which insurers assign their drug. The supplier can be expected to calculate the relative profitability of various combinations of F_B and co-pay waiver, based on the insurer's likely tiering and the supplier's estimate of the patients' elasticity of out-of-pocket demand.

Beyond that, the ability to adjust the patient's out of pocket payment turns co-pay insurance into another version of RP, albeit one which involves elaborate game playing between the supplier and the insurer. While a tiered co-pay structure might once have been seen as a device which could give the insurer greater control over its costs, the introduction of coupons seems to have transformed it into a high-transactions-cost version of RP. Further, the co-pay structure has the supplier picking the retail price, to which the insurer responds by setting an absolute amount co-pay, to which the supplier responds by determining the profit maximizing coupon amount. This game must be played between the insurer and each supplier who falls into a particular therapeutic category, since the objective of the exercise from the point of view of the insurer is to encourage substitution away from more expensive, and towards less expensive drugs. The RP system, on the other hand, lets the insurer specify the single per unit price that it will pay for any drug which falls into a particular therapeutic category. Both systems require the supplier to take account of the patient's elasticity of out-of-pocket demand, but it certainly seems likely that the transactions and information costs on the insurer's side will be less under RP than under tiered co-pay with coupons.

VI Policy Issues

The purpose of this paper has been to try and clarify the effects of three different, but common forms of insurance coverage on the prices of pharmaceuticals. Co-insurance, which was originally expected to hold prices down by tying the patient's out of pocket expenditure directly to the pharmaceutical's retail price level has proved lacking in that regard - in practice it permitted the per unit retail price of any quantity to rise until the consumer's maximum WTP for that quantity had been reached. Co-payment, being a blunter instrument, but which put the out-of-pocket price more directly under

the control of the insurer and allowed it to use tiered coinsured payments to encourage substitution in particular directions, had more promise but that promise has been blunted by the emergence of co-pay coupons funded, directly or indirectly, by pharmaceutical manufacturers. We have argued here that the effect of the emergence of such coupons was to turn tiered co-pay systems into versions of RP systems, but with higher transactions costs than a pure RP system would involve. Overall, it would seem that there is a strong argument for moving to RP on a wide front.

A number of policy questions arise with regard to the setting of the reference price. One is the issue of whether a country should use domestic or international reference pricing, the latter meaning that the reference price is set on the basis of prices in other countries. On the whole, while it is obviously desirable to be aware of what is going on in other countries, it seems preferable that one country's domestic reference prices be set on the basis of the value of domestic considerations - the domestic cost savings, for example, associated with switching from in-patient to pharmaceutical treatment as a result of the availability of new drugs. It also seems reasonable that higher income countries, which are likely to pay higher prices for a range of commodities, even traded goods, than are lower income countries, should be willing to pay higher prices for new entrant pharmaceuticals. The issue of the pricing of first-in-category pharmaceuticals is one which is unlikely to be easily resolved under any insurance form.

RP is clearly easier to implement when there are a number of drugs in a particular therapeutic category, whether all on-patent brand name drugs (with sufficient differences in their formulations not to violate each other's patents) or a mix of brand name and generic drugs. The presence of generics, however, raises another problem for RP - the reference price will become a floor, since there will be no particular reason for any supplier to set a price below the reference level. One reason the United States has had among the lowest-priced generic drugs in the world has been that competition among them has driven their prices down. This does not happen to anything like the same degree under RP systems.

The competitive mechanism has been so successful that commentators have often been taken by surprise when they discover that the market for generics is no different from the market for any other commodity - if a generic drug producer finds an opportunity to become a monopoly (without taking action that would violate anti-trust legislation) it will exercise its monopoly power and charge a monopoly price. It has also become evident recently that reliance on one or two manufacturers to supply a country's entire market for certain generic drugs means that the supply of essential drugs cannot be assumed to be secure. The departure of a manufacturer, or the closing of a firm's production facilities, even on a temporary basis, can create critical drug shortages. While there have been various proposals for regulatory remedies for such shortages, in the long term the best option is likely to be a combination of removing barriers to the importation of generic (not brand-name) drugs combined with the provision of an incentive to encourage supply. RP can do this if policy makers are willing to trade off higher generic prices against improved security of supply - keeping the reference price, even in the presence of a number of generic manufacturers, higher than the level which the free market would drive the price to (the free market limit price in practice would be roughly equal to the marginal cost of the physical production of drugs) in exchange for the greater security of supply which would follow from what would essentially be a guarantee of positive economic profit. In that case, should one generic supplier be forced out of a market, perhaps because of production problems, others would expand their operations to fill the gap, especially if

barriers to the importation of generics were lowered[12].

VII Conclusion

The desire to inculcate a degree of price sensitivity in patients who are having prescriptions filled has led insurers to develop a range of cost-sharing mechanisms. Each different mechanism creates a different structure on the demand side of the market, as perceived by suppliers of pharmaceuticals. While those suppliers are price setters, their ability to set prices is necessarily constrained by the demand side of the market, and they will choose their optimal price setting rules subject to the constraint which the demand side of the market imposes on them.

In this paper we have discussed, in graphical analysis, the optimal pricing behavior of suppliers of brand name pharmaceuticals faced with three common types of insurance structure - co-insurance, reference pricing, and co-payment. In addition, we have presented an analysis of the use of co-payment waiver coupons, a device which has been adopted relatively recently by suppliers, but which has increased rapidly in prevalence. We suggest that the use of co-pay waivers turns the co-payment insurance constraint into something very much like the reference pricing constraint, from the supplier's perspective, but with greater transactions costs. It may still be preferable, from the supplier's public relations perspective, however, since under reference pricing the supplier would be seen as adding a patient charge on top of the insurer's per-unit payment while co-pay waivers put the supplier in the position of cutting the price the patient might have had to pay.

Overall, we suggest that the supplier-pricing response to various types of insurance which a payer might be considering can be modeled reasonably well using standard economic tools, and that such modeling should precede any changes to insurance payment arrangements.

[12]We are thinking here in terms of reciprocity in the approval of generics and the inspection of production facilities among a number of countries whose drug regulatory systems are generally reliable.

References

Berndt, Ernst R., Thomas McGuire and Joseph P. Newhouse (2011): "A Primer on the Economics of Prescription Pharmaceutical Pricing in Health Insurance Markets" *Forum for Health Economics & Policy* 14(2)

Comanor, William S. and Stuart O. Schweitzer (2007): "Determinants of Drug Prices and Expenditures" *Managerial and Decision Economics* Special Issue: Economic and Policy Issues in the Pharmaceutical Industry, 28 (4-5), 357-370,June - August

Feldstein, Martin (1973): "The Welfare Loss of Excess Health Insurance" *Journal of Political Economy* 81(2), March-April, 251-280.

Freudenheim, Milt (2003): "Claritin's Price Falls, but Drug Costs More" New York Times May 8, 2003.

OECD (2015): Health at a Glance 2015: OECD Indicators, OECD Publishing, Paris
Pear, Robert (2016): "US To Test Ways To Cut Drug Prices in Medicare" New York Times March 8, 2016.

Robinson, James C. and Timothy T. Brown (2013): "Increases In Consumer Cost Sharing Redi rect Patient Volumes And Reduce Hospital Prices For Orthopedic Surgery" *Health Affairs* 32(8), August, 1392-1397.

Ross, Joseph S. and Aaron S. Kesselheim (2013): "Prescription Drug Coupons - No Such thing as a Free Lunch", *New England Journal of Medicine* 369, 13, September 26, 1188-1189.