
This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: http://openaccess.city.ac.uk/5359/

Link to published version:

Copyright and reuse: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.
Bundling in General Markets and in Health Care Systems

Xeni Dassiou
City University, London

Dionysius Glycopantis
City University, London

Charitini Stavropoulou
Imperial College Business School and the University of Surrey

Abstract. This paper provides an overview of bundling, by explaining, in general, the different forms that this practice can take in selling, buying or exchanging goods. We then proceed with a linear programming formulation in health care, to explain how the price of a bundle of resources can be calculated. We follow this by looking at the specific case of the practice of bundling by Medicare, the social insurance system for health care in the United States. We consider its advantages, disadvantages, and various recommendations.

JEL Classification: C61, D4, I11, I18
Keywords: Monopoly, Monopsony, Barter exchange, Price discrimination, Bundling, Pure bundling, Mixed bundling, Bundle price, Quality uncertainty, Adverse selection, Profit enhancing, Trade enhancing, Legal and regulatory systems, Linear programming, Primal and dual problems, Health care systems, Health insurance, Medicare.

1 Xeni Dassiou is a Senior Lecturer in the Department of Economics, City University, Northampton Square, London ECIV OHB, UK (e-mail: x.dassiou@city.ac.uk), Dionysius Glycopantis is Professor of Economics, City University, Northampton Square, London ECIV OHB, UK (e-mail: d.glycopantis@city.ac.uk) and C. Stavropoulou is Research Associate at the Imperial College Business School, South Kensington Campus, London SW7 2AZ, UK (e-mail: C.Stavropoulou@imperial.ac.uk) and Lecturer in Health Care Management at the School of Management, University of Surrey, Guilford, Surrey, GU2 7XH. All comments received on an earlier draft are greatly appreciated.
1. Introduction
The purpose of this paper is first to provide an overview of the theory, practice, and policy implications of the bundling of transactions in goods and services as an instrument of price setting and discrimination, as discussed within a recent set of papers. We look at some specific economic policy examples of bundling directly applicable to this theory. Second, we are interested in explaining through a simple mathematical model how the price of a bundle of resources of a hospital can be formed. We also discuss specific examples of bundling in the health care market which is of wide interest and importance.

The practice of bundling has increased rapidly, especially since the advent of e-commerce that has led into an explosion of trade in information goods and services. This has created the need for controlling the activity within a clearly defined legal and regulatory system.

The marked increase in bundling has lead to intensive research activity in the context of the theory of the firm. A number of models attempting to explain the economic principles involved and to predict the effects of bundling on profitability and the degree of competition has been set up. At the same time significant, from the point of view of their impact on the economy, real life applications of bundling have been considered by looking at the legal, regulatory, and competition framework within which this practice operates.

An early theoretical investigation of bundling was done by Adams and Yellen (1976) in which, as we repeat in some detail below, they also compare pure with mixed bundling from the point of view of extracting consumer surplus.

Bundling appears in many market forms. The practice of contract bundling by the U.S. government agencies is a good example of purchase bundling by a monopsonist. The Department of Defence in particular is the biggest bundler; it accounted for more than 65 percent of all bundling during the nineties.

The paper on monopsonistic mixed bundling by Dassiou and Glycopantis (2008) helps to explain why this may have been trade enhancing for the suppliers as well as profit enhancing for the government. In fact, this becomes increasingly more beneficial for the trade volumes enjoyed by the former if it is not easy to ascertain the quality of the all the goods involved in the bundle. We refer to this issue again later in this section.

We would like to explain that it falls outside the scope of this paper to review all the different proposed economic theories of bundling. This has been done by other authors, as for example by Kuhn (2004), who also discusses the role of bundling as an instrument of entry deterrence or product differentiation in the presence of competition. Rather, we use the papers by Dassiou et al. (2004) in which aspects of mixed bundling are discussed. In these papers bundling is seen as an

\footnote{Dassiou and Glycopantis (2005, 2006, 2008) and Dassiou et al. (2004).}
instrument of price discrimination within the context of various policy applications.

An innovative feature of our discussion is the derivation of the price of a bundle through a linear programming formulation of the problem of a hospital, with fixed quantities of resources providing health care. The solution of the dual problem gives a justification of the price asked of a health insurance company which is interested in buying the services of the resources as a bundle.

A discussion of the practical implementation of bundling in the context of health care is also provided. We concentrate on bundling by Medicare, and we discuss advantages, disadvantages, and various recommendations.

The paper is organized as follows: Section 2 discusses some aspects of bundling and provides an overview of the different forms and types that it can take, by clarifying that bundling can encompass sell and buy, as well as both types of actions in the case of countertrade. Section 3 discusses how the price of a bundle can be obtained in the context of a linear programming problem, and Section 4 considers issues of bundling by Medicare, the U.S. governmental agency. Section 5 concludes the discussion with some remarks.

2. Aspects of the Theory of Bundling and Its Practice
The typical form of bundling that the majority of the literature on bundling analyses, and can also be found in the majority of IO textbooks, is the case of a monopolist who bundles his goods to induce the buyers to buy the package. A familiar example is the bundling of articles in scientific journals, or, in turn, the bundling of journal subscriptions by publishers with the purpose of capturing the mean valuations of academics for a package of articles and journals; this facilitates the capturing of consumer surplus by the publisher since the bundling will reduce the dispersion in valuations unless the latter are perfectly correlated.

Bundling of services as set by the seller is also taking place in the case of airline travelling as this may includes inboard food, checked bags, check-in, and handling costs. More recently, the growth of internet flight booking has allowed many airlines to include in the package ancillary services such as hotels, car rental, and travel insurance booking, as they earn a commission from selling such services by third-party providers. On the other hand, it is important to point out that instances of unbundling can also occur as in the case of low-cost carriers. For example, Ryanair has recently unbundled its services by charging separately for inboard food, checked bags, and has more recently considered charging passengers for using the toilets on board its aircraft! It also plans to abolish facilities at airports by forcing all passengers to check-in on line and print off their own boarding passes before getting to the airport.

Unbundling of activities is also occurring in regulated utility industries. In the EU market for gas and electricity, there is a requirement of vertically unbundling the ownership or the operations of transmission networks owned by generation or energy supply companies. This has long been an unresolved issued despite numerous EU directives pointing to this direction over the last ten years. Such unbundling is seen
as a vital step for the development of competition within EU energy markets, and was only agreed to in 2008 by European ministers. Firms will be required to either sell their transmission networks or lease them to independent operators.

Another example of bundling is the practice by companies such as British Sky Broadcasting, Virgin Media, and more recently Carphone Warehouse, to sell bundles of services including phone, broadband and television services. (See Thanassoulis, 2008, on triple-play offers and its implications for consumers). BT received a boost from the regulator's recent decision (March 2009) to allow it to also offer such a bundle, whereas before, although it could sell such services separately, it was not allowed to market them as a single-bundle price item.

What all of the above examples share in common is that we have a price-setting seller offering the bundle. However, as argued in Dassiou et al. (2004) bundling does not only take the form of selling by a price-setting firm. It can also take the form of countertrade which can be interpreted as transaction bundling of one sell and one buy transaction between two partners, typically between a firm in a developed country (DC) and a trading partner in a less developed country (LDC).

In particular, while barter exchange can alleviate or resolve problems of creditworthiness from the side of the LDC (Marin and Schnitzer, 2002) and buyback contracts resolve problems of quality uncertainty in the case of technological transfers from a firm in a DC to a company in a LDC (Choi and Maldoom, 1992), countertrade can also act as a form of transaction bundling that allows the Western firm to price discriminate between potential trading partners. At the same time, it alleviates adverse selection regarding the uncertain quality of the good sold by the LDC firm.

By offering the opportunity to bundle its transactions, the DC firm can effectively enhance its ability to identify desirable trading partners and also engage in price discrimination. More importantly, barter does appear as a special case in exchange bundling. Dassiou et al. (2004) set the conditions under which offering this is optimal for the DC firm.

Finally, many procurement decisions take the form of bundling purchases. The firm will be willing to offer higher prices when bundling, and even more so if this allows it to resolve an adverse selection problem, if quality in any of the goods is uncertain. Just as it is profitable for a monopolist to offer mixed bundling at a bundle price which is lower than the sum of the individual prices (hence exploiting the 'average willingness to pay'), it is equally profitable for a monopsonist to offer a bundled purchase price which is higher than the individual prices being offered (hence exploiting the 'average willingness to sell').

Senate bills have been passed to control the practice of bundling in government procurement, while the Financial Services Authority and its counterpart across the Atlantic are at the final stages of forming policies that will regulate bundling in brokerage services and control the activity to the benefit of the society.

Another example is a health insurance company. Through its buying power over hospitals, it obtains the ability to control the physician fees despite the fact that
it lacks monopsony power over the physicians themselves directly, as argued in Blackstone and Fuhr (1996).

An early model of bundling is described in Adams and Yellen (1976). They define three different pricing strategies that a firm with monopoly power can pursue: (i) pricing and selling goods separately (pure components pricing); (ii) offering them as a bundle only (pure bundling); (iii) mixed bundling by selling the goods both individually and as a bundle. They identify as the chief defect of pure bundling its inability to comply with what they refer to as the condition of exclusion according to which no individual is allowed to consume a good which he or she values less than its marginal cost.

On the other hand, they argue that while mixed bundling still faces a tradeoff between more extraction of consumer surplus and more complete exclusion, it does so to a much smaller degree than pure bundling does. The use of mixed bundling as an effective mechanism for the extraction of consumer surplus is examined in a number of papers such as Stigler (1968), Adams and Yellen (1976), Schmalensee (1984), McAfee et al. (1989) and Salinger (1995).

Mixed bundling allows the firm to set individual component prices in a way that buyers, who value one of the two goods below its cost, cease to consume it while they continue to consume the other good which they value above cost. By setting the individual component prices properly, this practise ensures that mixed bundling is at least as profitable as pure bundling (Adams and Yellen, p. 483, footnote 12). On the other hand, if marginal costs are high, then pure components pricing will be more profitable than mixed bundling, as the cost of excluding customers in mixed bundling becomes too costly in terms of forgone consumer surplus.

Apart from the articles mentioned earlier on the use of mixed bundling, there has also been a number of articles on the bundling of information goods, first discussed by Varian (1985) and subsequently in a series of contributions by Bakos and Brynjolfsson (1998, 1999, 2000). These focus on the bundling of goods with low marginal costs. Such goods make mixed bundling a much more profitable strategy for firms relative to individual pricing as the possible inefficient consumption by customers who value one of the component goods less than its marginal cost is drastically reduced. In other words, satisfying the Adams and Yellen condition of exclusion is no longer an issue.

If the motive behind bundling is price discrimination, then this strategy is optimal for the firm independently of the existence of complementarities between the goods in the bundle (as discussed in Ausubel and Milgrom, 2002, and Lucking-Reiley, 2000). If anything, the existence of the latter would reinforce this result by creating an additional motive for bundling.

The profit-enhancing impact of bundling in the case where the goods bundled have independent valuations is examined by McAfee et al. (1989). They note that when the reservation values of the two goods are independently distributed, mixed bundling is always locally optimal relative to pure components pricing.
Bundling has already been shown to be profit enhancing by Schmalensee (1984) by reducing the effective dispersion of the reservation prices in a Gaussian demand environment, based on the fact that standard deviations of bundles are always subadditive unless the goods in the bundle are perfectly correlated. However bundling in that paper is based on the conventional combination of one unit of each good.

By Schmalensee's admission, if symmetry (in terms of values of the dispersion of the two goods by each being equal to half of the sum of the dispersion) does not hold, bundling is less likely to be profit- or welfare-enhancing. Hence, if the weight that each of the two goods receives in the bundle can be manipulated, then this may effectively resolve this problem without having to resort to a mixed bundling approach.

Moreover, a reduction in dispersion is not always profitable, but rather when marginal costs are low relative to the mean valuation. On the other hand, when marginal costs are high, the seller will want to increase, rather than decrease the dispersion of valuations because, if the marginal costs are greater than the mean valuation, bundling will decrease profits by decreasing the fraction of buyers with a mean valuation in excess of the marginal cost of the bundle.

For the case of a positive correlation, Nalebuff (2004) points out the superiority of pure bundling relative to mixed bundling as an instrument of entry deterrence. In mixed bundling the monopolist will have to charge very high individual prices to avoid having an emerging rival with one product to use the incumbent's other product to create a rival bundle that will steal away the established firm's bundled sales. But since such high prices mean that individual items will generate few additional sales, there is little point to offer mixed bundling in the first place. Hence we would not normally expect to observe mixed bundling in the presence of an entry threat.

A rather weak point in the above reasoning is that Nalebuff argues that the price discrimination effect of pure bundling results in limited increases in profits by offering a discount relative to the original component prices, while the main gain comes from the magnitude of the reduction in the potential profits of an entrant, thereby increasing the probability of entry deterrence. However, the modest profit increase that he refers to is that of pure bundling relative to pure component pricing. This ignores the fact that the profit increase can be much more substantial if there is a move from pure component pricing into mixed bundling. The difference in the magnitude between these two different types of profit gains needs to be considered, that is, the cost of choosing pure rather than mixed bundling, and then be traded off against the gains from entry deterrence accruing from pure bundling.

In other words, the distortion of bundling for strategic purposes will lead to anti-competitive effects and also forgone profits resulting from the choice of pure over mixed bundling. For this reason we argue in Dassiou and Glycopantis (2006) that regulation and competition authorities should focus on trying to stop bundling being used as an

---

3 In other words, unless the correlation in the reservation values between the demand for the two goods is equal to 1 (\(\rho = 1\)), \(\sigma_B \leq \sigma_1 + \sigma_2\).
instrument of entry deterrence, rather than trying to ban bundling altogether.

As mentioned, the novelty of the papers by Dassiou et al. (2004) has been that rather than looking at bundling in terms of a monopolist arrangement only, we consider three distinct possibilities. We analyze transactions bundling in the cases of monopoly, monopsony, and exchange (organized by a price-setting firm which is both a monopolist and a monopsonist). The common result is the local optimality of the bundling of transactions from the point of view of expected profits for the price-setting firm, and the overall increase in the level of trade in the goods bundled.

Contrary to the analysis by Nalebuff (2004) we assume no complementarity between the goods bundled. This removes the possibility of the use of bundling for strategic purposes, i.e., as an instrument of foreclosure, retaining the use of bundling for short-run profit maximizing purposes alone, as practised by an uncontested price-setting company.

What is considered instead, for the cases of exchange and monopsonistic bundling (Dassiou et al., 2004, Dassiou and Glycopantis, 2008), is the issue of varying degrees of adverse selection (what we refer to as partner preference), and more specifically the benefits arising from bundling not only in terms of the increase in the profits of the firm that sets the prices, but also in terms of enhancing the volume of trade in goods with some degree of uncertain quality.

The reason why the enhancement in the volume of trade emerges in these two papers as a separate issue is obvious. In the standard case of mixed bundling in monopoly, the resulting increase in the volume of trade translates into an increase in the profits of the price-setting firm. In the case of exchange and monopsony, gains in the volume of trade are also experienced by the trading partners of the price-setting company.

While Nalebuff (2004) and Salinger (1995) discuss bundling in terms of creating a more valuable product by tying two complementary goods as one, the Dassiou et al. (2004) and Dassiou and Glycopantis (2008) papers, by looking at exchange and monopsony respectively, show how the bundling of transactions helps price takers to resolve the problem of quality uncertainty in the good(s) that they sell. Bundling, as we see, increases the volume of trade; this increase is enhanced considerably if the purchase of a good with low quality certainty is bundled with another transaction, buy or sell, of a good with a substantially higher degree of quality certainty.

To the best of our knowledge, the only other paper which looks at bundling a product of established quality with one of unknown quality as a means of mitigating the problem of asymmetric information is the one by Choi (2003). He interprets bundling as a way to leverage information and signal quality, and finds that it is profitable when the quality of the one product is known while the quality of the other is unknown.

The paper on exchange bundling (Dassiou et al., 2004) concurs with Choi’s finding; the difference is that, in the two, products are sold by two different

---

4 Carlton and Waldam (2002) examine the use of pure bundling of complementary goods, in the form of tying, in order to prevent entry into the market.
companies. The seller of the good of certain quality is the price-setting monopolist in a developed country. He mix-bundles this sale with the monopsonistic purchase of a good of uncertain quality produced by a company in a less developed country. We find that bundling of a (sell, buy) transaction is profit-enhancing, unless either the marginal cost of producing the good sold is extremely high (i.e., equal to one), or if the certainty component in the quality of the good bought is zero.

Choi argues that bundling is never profitable if consumers either know the quality of both goods or they are not certain of the quality of either. The monopsony paper by Dassiou and Glycopantis (2008) concurs this latter finding but refutes the former. The finding by Choi is not common in the literature. It is the result of the restrictive assumption that the monopolist faces identical customers. Dassiou and Glycopantis (2008), on the other hand, in common with the majority of other authors, assume different types of trading partners. Hence mixed bundling in our model is superior, even if both goods are of certain quality, as it reduces the dispersion in the valuations (unobservable by the price-setter) of the sellers. In general, it is optimal for the price-setting monopolist to mix-bundle the purchase of two goods together, (relative to no bundling), unless the value of the certainty component in the quality of either or both goods is zero. The difference is that Choi talks about a conventional model of bundling by a selling firm producing two goods.

He shows that the bundling of a good whose quality information is imperfect with a product whose quality can be fully ascertained can act as a signal of quality for the former good. Interestingly he also argues that this practice of pure bundling may seriously backfire if the weakest-link good degrades the performance of the high-quality good when consumed together. As an example of this, he discusses Merill Lynch's experience with its Cash Management Account.

This account was basically a bundling of financial services by permitting card and check withdrawals from an investment account. It was therefore an attempt to enter retail banking by bundling the characteristics of a current account to those of an investment account. Merill Lynch encountered significant operational problems with its account because of the time-consuming customer enquiries directed to its brokers. The company eventually had to hire several hundred clerks for the programme. In essence the adding of this cheque and card facilities degraded, rather than enhanced, the value of its investment account.

On the other hand, Dassiou and Glycopantis (2008) look instead at a monopsonist's arranged mixed bundle and deduce that the larger the difference in the degree of quality uncertainty between the two goods is, the larger will be the improvement in the volume of trade for the good with the higher degree of quality uncertainty relative to no bundling. At the extreme case where the quality of one of the goods bundled can be fully ascertained, the increase in the volume of trade for the other one is maximized. This maximum value is larger the higher the degree of quality uncertainty in this second good. Hence, packaging the transactions of two goods that are independent of each other and of different qualities induces an increase in the volume of trade for both goods and especially in the one with the high
degree of quality uncertainty.

Furthermore, for the price-setting monopsonist there will be an improvement in profits relative to no bundling unless, as in Choi, the degree of quality certainty in either or both goods is equal to zero. This improvement in expected profits will be greater, the smaller the difference in the degree of quality uncertainty between the two goods purchased is.

DeGraba (2005) deals with a risk-averse monopolist who prices closer to marginal cost when demand is uncertain in order to secure this demand as it is considered riskier that many small buyers. Dependence on a larger buyer creates higher uncertainty, and hence the risk-averse seller seeks to avoid this by offering a lower price to such a large buyer than the price demanded from smaller buyers. Obviously the relevance of such a paper in the context of bundling is two-fold: firstly, it offers an explanation for the tendency by the government to bundle its purchases as an attempt to improve its bargaining power and secure a better deal from a price-setting monopolist, and secondly it explains why the seller wishes to offer the bundle at a discount other than the price discrimination reason.

Crampes and Hollander (2005) look at bundles in pay-TV programme content provision and find that a firm is likely to offer two bundles in order to increase profits when the majority of preferences over one type of programming is strong and when the gap in consumer preferences is large. When two distinct bundles are offered to subscribers, only one of them will contain all channels while the other will be a subset of the former. Hence we have the typical case where there is a basic service open to all subscribers, and then additional channels are offered to those willing to pay a premium. This achieves screening, by versioning, and allows the full extraction of the consumer surplus from those with the highest willingness to pay.

Essentially this paper introduces both the idea of weighting, (for example, how many type 1 (good 1) and how many type 2 (good 2) channels should be in the bundle), as well as the possibility of offering alternative bundles where some have a full coverage and some partial cover. We take up this second theme in the modelling section below, by considering different packages, (full and partial), of illnesses coverage in the case of hospitals. The weighting of the scarce resources comes from the solution of the dual problem.

3. An Application of Linear Programming
We discuss here how the price of a bundle can be formed through the solution of a mathematical model. The idea is that there is a hospital with fixed quantities of factors of production, doctors, nurses, beds, and medicine, and can accept patients who suffer from heart problems, cancer, stroke, or a common cold. The hospital can admit directly patients who pay a given price for a fixed period of treatment, or it can come to an agreement with a health insurance company which is prepared to pay a price for the bundle of all, or part of, the available resources. The insurance company makes a profit by collecting a premium from people who are risk averse and wish to
insure themselves against such an illness.

The hospital is faced, first, with the problem of maximizing total revenue subject to feasibility constraints.

**The primal problem:**
Maximize \( R = p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 \)
subject to:
\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 & \leq q_d \\
    a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 & \leq q_n \\
    a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 & \leq q_b \\
    a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 & \leq q_m \\
    x_1, x_2, x_3, x_4 & \geq 0.
\end{align*}
\]

**Interpretation of the primal; variables and relations:**
For simplicity we assume divisibility of the factors of production. In a particular period the hospital can admit:
\[
\begin{align*}
    x_1 & \text{ number of heart patients, each paying } p_1, \\
    x_2 & \text{ number of cancer patients, each paying } p_2, \\
    x_3 & \text{ number of stroke patients, each paying } p_3, \\
    x_4 & \text{ number of common cold patients, each paying } p_4.
\end{align*}
\]
The prices that will prevail and the number of patients to be attracted are in expectation form and therefore \( R \) denotes the expected profit. The factors of production available to the hospital during the same period are:
\[
\begin{align*}
    q_d & \text{ the quantity of doctor-hours available, (Resource 1)}, \\
    q_n & \text{ the quantity of nurse-hours available, (Resource 2)}, \\
    q_b & \text{ the quantity of bed-hours available, (Resource 3)}, \\
    q_m & \text{ the quantity of medicine available (Resource 4)}.
\end{align*}
\]
With respect to the technological coefficients, \( a_{ij} \) denotes the non-negative units of resource \( i \) required per unit of patient of type \( j \). Therefore the technological constraints express the fact that it is not possible to allocate more than the available quantities of the factors of production. We also have the feasibility constraint that the number of patients of each type must be non-negative. Of course, some of the \( a_{ij} \)'s could well be zero. For example, it is possible that the treatment of common cold does not require that the patient stay in a hospital bed for any period. The set of constraints is assumed to define a feasible set, and therefore a solution exists. It is
denoted by \((\mathbf{R}^*, x_1^*, x_2^*, x_3^*, x_4^*)\), patients of each type that the hospital would like to attract ideally, and the revenue that it would like to make. The formulation above is the usual one in economics for finding the optimum of an objective function subject to feasibility constraints. Problems in linear programming come in pairs and every one has its dual. The formulation of the latter uses the variables attached to the constraint of the primal. This is well known.

**The dual problem:**

Minimize \( C = r_1q_d + r_2q_a + r_3q_b + r_4q_m \)

subject to:

\[
\begin{align*}
   a_{11}r_1^* + a_{21}r_2^* + a_{31}r_3^* + a_{41}r_4^* & \geq p_1 \\
   a_{12}r_1^* + a_{22}r_2^* + a_{32}r_3^* + a_{42}r_4^* & \geq p_2 \\
   a_{13}r_1^* + a_{23}r_2^* + a_{33}r_3^* + a_{43}r_4^* & \geq p_3 \\
   a_{14}r_1^* + a_{24}r_2^* + a_{34}r_3^* + a_{44}r_4^* & \geq p_4
\end{align*}
\]

\( r_1^*, r_2^*, r_3^*, r_4^* \geq 0 \).

**Interpretation of the dual; variables and relations:**

The economic interpretation of the dual is, of course, related to the physical problem described in the primal, especially since their optimal values must be identical. Here, we create a shadow market for the factors of production where \( r_i \) is the required, non-negative, accounting price of resource \( i \), i.e., its weight, and we mimic the competitive solution. The objective function expresses the fact that we wish to minimize the total cost of hiring the fixed factors of production in the shadow market by observing at the same time the competitive equilibrium conditions.

Each of the constraints refers to a particular type of patient and expresses the requirement that in the production of the corresponding services marginal cost must be greater or equal to marginal revenue. Marginal cost for each type of patients is expressed by summing up the factor requirements times the corresponding shadow price and marginal revenue by the price \( p_i \). In the case where marginal cost is greater than marginal revenue the corresponding type of patient is not admitted in the hospital as this would be very expensive. These are the classical competitive equilibrium first order conditions.

We are trying to allocate the total expected profit to the scarcity of the factors of production. Let the solution of the dual be \((C^*, r_1^*, r_2^*, r_3^*, r_4^*)\). We have \( r_i^* = \frac{\partial R^*}{\partial q_i} \)

where \( i \in \{d, n, b, m\} \). Each partial derivative denotes the per unit contribution of the respective factor of production to the maximum \( R = R^* \). Therefore this is a signal as to what resource new investment should be directed to.

The duality theorem implies that \( R^* = C^* \). Based on this fact, we can create a
bundle of resources at the maximum price $R^*$ or $C^*$. It is precisely the solution of the dual problem that gives the competitive value, as calculated by the hospital, of the bundle of resources. Below we consider a number of negotiating circumstances between the hospital and the insurance company.

**Case 1.** The Hospital and the Insurance Company negotiate over the whole package of resources.

Everything is in expectation. Therefore, on the one hand, the hospital can profit from securing the expected profit, $R^*$, in advance, and the insurance company can argue about the amount of money demanded for the whole bundle of resources. The insurance company can profit if there is a sufficient number of risk averse individuals prepared to pay a sufficiently high premium for complete cover. For simplicity we assume there is only one type of such individuals.

Figure 1 shows that a risk-averse individual is prepared to pay a premium which is larger than the expected loss. $W_0$ denotes wealth at the current state of health and $W_0 - \hat{I}$ the level following an illness. $E(L)$ is the expected loss in wealth, due to the probable illness, and $G$ the insurance premium that a risk-averse individual is prepared to pay to obtain coverage. The fact that $G > E(L)$ allows the insurance company to operate and make a profit. As the individual insures himself against a number of alternative illnesses, we can assume that $E(L)$ is an average expected loss which can occur with a given probability.

The insurance company is asking the hospital to accept, for the whole bundle, a lower amount $r^*$. This is based on its own estimation of future prices and profit. We suppose that the hospital considers $r^*$ sufficient, if not optimal, to justify its operations. Therefore there is an amount $R^* - r^*$ to be divided, and this depends on the bargaining power of the two sides. Figure 1 depicts the case where the middleman between the hospital and the patient has more bargaining power.

One way to proceed is to assume that bargaining takes place between the two parties. We can, for example, invoke from game theory the axiomatic Nash bargaining solution, (Nash, 1950), of the problem below:

**Problem:**

Maximize $U = y_1^\alpha y_2^\beta$

subject to:

\[ y_1 + y_2 \leq R^* - r^* \]

\[ y_1, y_2 \geq 0, \]

where $y_1$ is the amount of the difference between $R^* - r^*$ to be allocated to the insurance company, $y_2$ to the hospital, and $\alpha, \beta \geq 0$, summing up to one, denote the
Figure 1: A Risk-averse Individual: Premium > Excepted loss

Figure 2 presents the case where the hospital has less bargaining power. The solution gives the bundle price, $r^* + y_1^*$, that is agreed on for the resources. The insurance company has avoided paying an extra $y_1^*$ for securing the bundle of resources.

We note that in the above formulation we have taken the status-quo payoffs to be equal to zero. This assumption means that if negotiations break down, both the hospital and the insurance receive zero income. It would be more realistic to assume that the hospital still makes a positive income $S'$, which will be its status-quo payoff. This would be lower than $R'$ because the prices will be affected from the fact that negotiation were entered into unsuccessfully. The mathematical formulation of the problem would be, of course, easy to adjust.

Alternatively, without invoking the strict axiomatic approach of Nash, we can divide $R' - r'$ between the two parties according to some agreed to proportions expressing the behavioural attitudes of the two parties in desiring to settle and to make a profit. A bounded rationality argument can determine these proportions, without the need to invoke a set of axioms to support the decisions of the agents.
Case 2. The Hospital and the Insurance Company negotiate on payment over a percentage of the whole package of resources.

This could correspond to the case where the hospital is not just a profit pursuing organization, but it also wishes to observe social norms and serve the community. In particular, the exclusion from bundling of cases involving patients with severe conditions may remove one of the disadvantages of bundling in health care discussed below. Such exemptions, combined with the setting of targets on health outcomes, will reduce the tendency to economize by avoiding accepting very serious, and therefore, costly cases. On the other hand, the insurance company might have established through market research that there is a substantial number of people who are optimistic about their future state of health, ready to take a risk. Therefore, they would not be prepared to pay a reasonable premium for health insurance. It might also be the case that, although risk-averse, the market is not big enough to bear the burden of securing the payment to the hospital and making a reasonable profit.

A way forward is as follows: Let \( \lambda \) be the percentage of the bundle of resources under negotiation. The hospital solves again the primal problem and asks for a payment of \( \lambda R^* \), the insurance company offers \( \lambda r^* \) and we are faced, in terms of negotiations, with a reduced version of the negotiation problem of Case 1.

Case 3. The Hospital and the Insurance Company negotiate over part of the package of resources.

We wish to discuss the case where the hospital is not supplying all types of
medical activities. This occurs when, for example, the insurance company does not want to cover all people who might catch a common cold. In the opposite case, the premium might have to be reduced because most people will probably believe that they will simply catch a cold. Also, the hospital might wish, or be expected by the National Health Service, which would be an external constraint on the problem, to treat this type of patient. In this case the hospital solves again the primal linear programming problem above and asks from the insurance company the amount \( P^\prime = p_1 x_1^\prime + p_2 x_2^\prime + p_3 x_3^\prime \). The latter offers \( P^\prime \) and we have in a reduced form the problem of Case 1. The original primal problem is solved because if the hospital were to withhold resources and then solve the sub-problem that would emerge, it could not obtain a larger solution than \( R^\prime \).

We cannot do any better by dividing the problem into smaller linear programming sub-problems. The hospital could not obtain a larger solution than \( R^\prime \) because in the end the overall constraint on the resources must be satisfied. Of course, it must take into account that not all of \( P^\prime \) will be secured. On the other hand, all sums are in expectation, and what is described here is one way of proceeding.

In the discussion above, we have used an example from health economics, in the form a linear programming problem, to explain how the price of a bundle can be determined. Obviously, this type of problems can also be formulated in the context of other areas in economics, and analogous results concerning bundles and their prices can be obtained.

4. Applications from Health Economics and Health Care Systems; Medicare

We shall now discuss bundling in the health-care market. Because of the relative availability of sources, we concentrate on Medicare, the insurance system in the United States which offers coverage to certain groups of people.\(^5\) It acts as a buyer, with monopsonistic powers, of the services provided by hospitals.

Bundling in health care:

In every health-care system, insurance plays a significant role. Since Arrow’s (1963) pioneering work on the health-care market, a systematic effort has been made to understand the economic and behavioural factors which drive the dynamics of health insurance. Health insurance systems around the world are faced with a number of challenges, including the ageing of the population, new medical technologies, and increased patients’ expectations (Waters et al., 2008).

The bundling of care has been discussed as a way of improving efficiency in health services. By bundling payment in the context of health care we mean payment by episode of care, rather than payment for individual services. For example, an

\(^5\) The provision of health care coverage in an economy is important for the population longevity and general quality of life, (see Sen, 2009).
insurance package pays a fixed amount to hospitals for knee replacements which include, as a bundle, both acute and post-acute care, instead of paying separately for hospital stay, doctors, nurses, and medication needed.

The U.S. health system, in particular, has struggled for decades to maintain efficiency and fairness under Medicare, a social insurance programme administered by the government to provide health coverage to people who are aged 65 and over, or who meet other special criteria. This creates significant pressure to identify ways to decrease costs, increase efficiency and provide broader coverage of the insured population. The Medicare Payment Advisory Commission in its recent report to Congress discuss extensively bundling payment reforms as an effective way to promote the effectiveness and efficiency of Medicare in the United States (MedPAC, 2008). Yet, bundling in health care is a relative new approach in health insurance systems, and its implementation has been limited to only a few pilot schemes designed by Medicare.

Below we discuss a number of advantages and disadvantages that the debate on bundling has identified, as well as the lessons that have been learned from the implementation of bundling.

**Advantages of Bundled Health Care:**
One of the main advantages of bundling in the health care context is that bundling can eliminate unnecessary services and consequently reduce costs. Providers have an incentive to decrease the number of physician services that are not needed during hospitalisation, use them in a more appropriate way, and reduce post-acute care and readmissions. Indeed, it has been estimated that Medicare spends $12 billion on readmissions that could be prevented (MedPAC 2008).

The current fee-for-service (FFS) system pays providers a prospective amount for services based on the expected cost of these services, and therefore, by providing more services, such as more admissions, longer stays and more tests, some of which could be avoided, hospitals increase their income. For example, Medicare pays hospitals per length of stay, so that the cost of admission is the same with the cost of readmission. It therefore provides no incentive to the hospital to avoid readmissions.

While fee-for-service payments reward more care, a bundled payment method incentivizes providers to better allocation of care. The limited evidence on bundling payment in the United States confirms a number of positive effects that the programme has had on reduction of spending. Most evidence comes from the Health Care Financing Administration’s (HCFA) Medicare Participating Heart Bypass Center Demonstration. The programme included a bundling payment for coronary artery bypass graft at four hospitals in the United States and ran between 1991 and 1996.

The study by Cromwell et al. (1997) evaluating the programme finds a 15.5 percent decrease in Medicare spending between 1991-1993, saving $17.2 million to Medicare, of which the greatest proportion was in-patient savings. A study by Wynn (2001) shows that the savings from bundling totalled $52.3 million, of which $42.3 million was due to discounts negotiated with the hospitals and $7.9 million was
from reduced co-payments.

A more recent implementation of the bundling system in the U.S. health-care market was the “ProvenCareSM” programme implemented by the Geisinger Health System in 2006. ProvenCareSM included a bundling payment system for all non-emergency coronary artery bypass graft (CABG) procedures and was accompanied by a pay-for-performance process (Paulus et al. 2008). The bundle price included the estimated cost of a typical hospitalisation and ninety days of post-acute care.

A study by Casale et al. (2007) looking at the scheme shows that, indeed, the programme led to a better provider compliance to the guidelines and dropped hospital costs by 5 percent. It also shows that the length of stay, as well as the readmission rates, declined significantly. The success of the scheme led Geisinger Health System to expand bundling to hip-replacement surgery and also considering expansion in to knee replacement, cataract surgery and other conditions (Miller 2008).

In addition to reducing spending by controlling readmissions, providers have an incentive to reduce unnecessary physician services during the hospitalization or switch from doctors to nurses to reduce costs when appropriate. Also, physicians are driven to use fewer resources during an in-patient stay and use cheaper alternatives to therapies, for example, generic drugs instead of branded ones, or low-tech services instead of high-tech which are equally effective.

Finally, patients may benefit from sharing the savings resulting from reduced service utilization and costs (RAND, 2008). This advantage becomes more obvious when patients have to pay part of the cost of treatment.

**Disadvantages**

However, the bundled payment method is not without considerable limitations and potential problems. First, there is a danger that a bundling payment system will generate unwanted incentives to providers to increase their income or eliminate costs in undesirable ways. For example, given that providers are paid by episode of care, they may increase admissions to hospitals as a way of increasing income, even in situations when the patient could be treated on an outpatient basis. Also, along with reduction in unnecessary services to eliminate costs, there is the danger that hospitals may also reduce necessary services. This creates concerns regarding the impact of the system on health outcomes.

In general, the implementation of bundling payment methods should in theory improve health outcomes of patients. Providers have an incentive to coordinate better with each other, to use different resources and to reduce the probability of re-admission and unnecessary hospitalization. However, if the target is only cost-containment, there is the danger that providers may cut down on not only unnecessary care but also appropriate care. In this case health outcomes may be reduced. There is also the danger of selection bias if severity of the condition is not taken into consideration. Bundling payment is based on average costs, and hospitals
Evidence on the impact of bundling on health outcomes is limited and therefore does not allow for generalizations to be made. Casale et al. (2007) find that adverse effects declined among patients who were part of the bundling scheme in comparison to those patients who were treated a year prior to its implementation in 2005. Yet, patients in the study were admitted for elective CABGs and so tended to be healthier than the general population of patients with the same condition, and the sample was too small to allow generalization of results. To avoid unwanted consequences in health outcomes strict quality of care standards needs to be set by the authorities.

The impact of bundling on patient experience is not known. In theory, there may be a positive effect due to better coordination between providers, but there is concern that doctors may be forced to spend less time with their patients as a result of an effort to see more patients and reduce cost. This is likely to affect patients’ satisfaction levels. An additional limitation is that because of disparities in the financial performance among hospitals, some hospitals may be able to pay physicians higher rates than some others can. That may lead hospitals to redirect money from other sources, such as nursing, to physicians in order to offer attractive compensation arrangements. Other limitations concern implementation in a broader scale.

Bundling has been proven useful in reducing costs and increasing savings in a number of conditions, but it may be less easy to implement the system when the duration of the illness is less clearly defined, such as in chronic conditions (e.g., diabetes).

Operational changes are needed to implement bundling payment. The complexity of the episode care may make it difficult for physicians and hospitals to decide on the most appropriate way of managing the care and sharing the bundling payment (Berenson et al., 2006; Budetti et al., 2002). This is done ad hoc per case. A number of decisions need to be made, including which conditions can be part of the scheme, the severity of the patients considered for participation, ways to achieve the most efficient bundles and mechanisms than ensure that appropriate measure is taken to avoid the danger of reduced health outcomes. That involves the danger of conflicts of interest among the different parties, such as doctors, hospital managers, and commissioners. Similar conflicts may occur when the determination and distribution of payments is discussed.

Recommendations

The Medicare Payment Advisory Commission (MedPAC 2008), in view of the potential drawbacks that the implementation of a bundling system has, proposes a number of recommendations to make its implementations more effective. First, it suggests that information concerning service use around hospitalization episodes should be disclosed to hospitals and physicians, to allow them to compare their performances relative to that of others and to examine how well they are doing in

---

6 If the patients had to pay their cost of treatment that would have been a problem of adverse selection.
readmission rates. It suggests that this measure, which will help physicians and hospitals to reduce their service use, be first disclosed to hospitals and physicians only and then the information be made available to the public.

Second, it proposes that for the first measure to work, it should be implemented along with financial incentives. That would include Medicare reducing the payments to hospitals with relatively high readmission rates for selected conditions. This would also allow hospitals to share the savings from episode care treatment, by reducing the cost below the Medicare bundle payment.

Third, to fully understand the potential practical problems that might occur from the implementation of a bundling payment system, the Committee suggests the creation of a pilot programme to test the feasibility of the system, initially in a number of conditions. This study would identify the potential problems of the method and will suggest whether it is feasible to extend the system to more conditions under Medicare. To avoid the danger of reduced health outcomes due to reduction of appropriate care in order to decrease costs, it is important to set targets to maintain health outcomes. Also, the scheme needs to take into account the severity of the condition as sicker patients may be avoided for being financially riskier than others. Also, as it is easier for the bundling system to be implemented if the duration of an episode can be defined, but more careful design is needed for conditions where the starting and ending points are less clear.

More pilot studies are needed to examine the practical problems occurring from the implementation of bundling systems. So far, the experience from the U.S. implementation of bundling systems gives encouraging results for the reduction of costs and spending, and a wider implementation would provide more evidence for the scheme's potential.

5. Concluding Remarks
We have discussed a number of issues. A typology of bundling should involve more than the distinction between mixed and pure bundling. The literature has focused on bundling by a selling firm which can set a bundle price as well as separate prices. However bundling can also take alternative forms which depend on how the bundle is formed. Two such forms are “bundle setting” by the buyer, as in the case of procurement, as well as a “bundled exchange”, i.e., countertrade. Once bundling is considered in this light, then its implications have to be examined not only in terms of profit for the price setter, but also from the point of view of its effects on trading for the firms transacting with it. Furthermore, apart from some exceptions, the possibility of adverse selection, which has only received very limited coverage in the literature, introduces the need to examine the implications of bundling in alleviating questions of partner preferences. These could stem from quality uncertainties and hence further encourage trade, through the pairing of goods with different qualities. Hence such uncertainties introduce another potentially welfare-enhancing dimension of bundling. Finally, the other common perception typically found in the literature is the equal weighting of all goods bundled in the basket. Of course, this is not
necessarily the case in all instances. For example, the weights of types of programmes, such as documentaries, sports, movies, news, etc., bundled in the offerings by a TV provider can vary. This can lead to the question of how the firm will determine the weights of the goods within a bundle that it sells or wishes to buy, and the welfare implications of such decisions.

The linear programming formulation discusses an approach that can be taken to obtain the price of a bundle of medical resources and the possible tension between a hospital and an insurance company. The theoretical discussion of bundling is followed by an application in the health care context. In particular, we discuss the implementation of bundling payments by Medicare, a social insurance programme in the United States. A number of pilot implementations of bundling in coronary artery bypass graft show positive effects in terms of reducing readmissions and health-care spending. However, these studies reveal also a number of drawbacks. There is a danger that hospitals will increase admissions in order to increase profit. They may also avoid sicker patients who increase costs. Furthermore, it may be difficult for bundling to work for chronic conditions where the starting and end points of care are less clear. In order to avoid unwanted consequences, strict quality targets need to be set and more pilot studies to be implemented.

6. References


