Health Economics for Health Professionals
An Aotearoa / New Zealand Perspective

Guy Scott
Table of Contents

Foreword ........................................................................................................................6

Acknowledgements ........................................................................................................6

Chapter 1: Basic Economic Concepts of Relevance to Health Care .............................7

Objectives and Key Concepts ........................................................................................7
Definitions of economics and health economics ...........................................................7
Economic resources .......................................................................................................7
Basic economic questions of scarcity, choice and rationing ..........................................8
Adam Smith: the invisible hand and the free market.....................................................8
The production possibilities frontier (PPF) .................................................................9
The law of diminishing returns ...................................................................................12
Economies of scale ......................................................................................................12
Efficiency (technical or productive and allocative) .....................................................12
Equity ...........................................................................................................................12
The nature of health and the determinants of health status ........................................13
Questions to consider ...................................................................................................14

Chapter 2: The Market for Health Care .......................................................................16

Objectives and Key Concepts ......................................................................................16
Demand ........................................................................................................................16
Utility and marginal utility ...........................................................................................16
The law of demand .......................................................................................................16
Reasons why the quantity demanded reduces as a good’s own price increases ..17
Describing demand .....................................................................................................17
Determinants of demand ............................................................................................17
The difference between a shift in demand and a change in quantity demanded..18
Demand, benefits and willingness to pay ..................................................................19
Supply ..........................................................................................................................19
Supply, sellers and cost ................................................................................................19
Law of supply ..............................................................................................................19
The interaction of market forces and price rationing ...................................................21
Questions to consider ...................................................................................................23

Chapter 3: Elasticities of Demand and Special Features of the Market for Health
Services ........................................................................................................................25

Objectives and key concepts .......................................................................................25
Elasticities of demand ...................................................................................................25
Own price elasticity of demand (PED) .........................................................................25
Income elasticity of demand (YED) .............................................................................27
Cross price elasticity of demand (XED) .......................................................................28
Chapter 7: Costs and Effects of a Health Care Intervention or Policy ...............51

Objectives and key concepts.................................................................51
Costs and Money..................................................................................51
Opportunity Cost..................................................................................51
Perspective and differences between decision-makers................................52
  Internal or private costs and benefits..................................................53
  Social (or public) costs and benefits....................................................53
  Externalities or spillovers (external costs and benefits)......................54
  Examples of external costs and cost shifting/ transfer costs................55
  External benefits ...............................................................................55
Cost classifications in health care.........................................................56
  Direct costs ......................................................................................57
  Indirect and intangible costs ..............................................................58
  Ease of identification and measurement ............................................58
Health related effects...........................................................................58
  Qualitative phase in estimating costs and effects...............................59
  Quantitative phase in estimating costs and effects..............................59
Questions to consider............................................................................60

Chapter 8: Economic Evaluation in Health ...........................................61

Objectives and key points .................................................................61
Theoretical foundations of economic evaluation.....................................61
The health economic evaluation process ..............................................61
Types of economic evaluation in health care.........................................62
Decision-making..................................................................................63
Health economic modelling..................................................................65
Check lists for assessing economic evaluations....................................65
The economic evaluation process.........................................................65

Books and Websites you may find useful.............................................68

References..........................................................................................68

The author ..........................................................................................73

Figure 1. The production possibilities curve depicting education services versus health services .................................................................10
Figure 2. A production possibilities curve illustrating increasing opportunity cost ..........................11
Figure 3. A production possibilities curve illustrating constant opportunity cost .......11
Figure 4. Some health related quality of life dimensions ..........................13
Figure 5. Determinants of health status ................................................14
Figure 6. Movement along and a shift in demand ....................................19
Figure 7. The market for physiotherapy consultations ..............................22
Figure 8. The Grossman model.............................................................28
Figure 9. Bradshaw’s categories of need for health care ..........................29
Figure 10. The demand for health care ..................................................30
Foreword

This monograph draws upon lecture notes created for a post graduate health economics course at Massey University New Zealand. The focus of the text is on the practical application of economic concepts to health. The theoretical concepts and examples provided are of particular relevance to a range of health professionals and to the delivery of health services in mixed market economies. The monograph is intended to be a summary of economic concepts relevant to the health sector, it is not intended to replace more detailed and theoretical health economic texts or journal articles.

Some specific examples of the types of issues that the application of health economics could help resolve are as follows:

- What are the roles of the market and government with respect to improving allocative efficiency and social equity?
- How can health care resources be allocated to achieve enhanced health outcomes for a defined population?
- How much should New Zealand spend on health?
- What health services should the health sector deliver?
- How can we improve the efficiency of healthcare delivery?
- How do we set priorities in health care provision and delivery?
- Who should receive healthcare services?
- How can we use the resources devoted to health to best improve the nation’s health?

Acknowledgements

I wish to thank Helen Scott and Terry Auld for their assistance in peer reviewing the text and the librarians at the Massey University for their advice on publication.
Chapter 1: Basic Economic Concepts of Relevance to Health Care

Objectives and Key Concepts

After reading this chapter you should:

- Appreciate the key economic questions of what to produce, how to produce it and who will get the goods produced.
- Be able to identify different types of health care resources.
- Be familiar with the concepts of scarcity, choice and opportunity cost.
- Understand how the production possibilities frontier may be used to illustrate the concepts of scarcity, opportunity cost and choice.
- Realize the difference between efficiency and equity.
- Appreciate the role of health economics in the efficient and equitable provision of health care.
- Appreciate some important basic economic ideas relevant to health economics.
- Be aware of the nature and determinants of health.

Definitions of economics and health economics

Economics is a social science concerned with determining how to best use society's limited resources to satisfy human desires. Health economics is a branch of economics that addresses the allocation of scarce resources to improve the health status of individuals and populations. Other health economists’ definitions may be worded slightly differently.

Economic resources

Resources (factors of production or inputs) may be used to produce desired goods and services. Historically, resources have been classified as land, labour or capital. The meanings of land and of capital, as defined by the early economists is as follows.

*Land* or natural resources includes all naturally occurring resources such as unimproved land, indigenous forests, minerals, fisheries, water and wind.

*Labour* includes both physical and mental human effort. Enterprise (the managerial and risk taking component of mental effort) is sometimes recorded as a separate category.

*Capital* (or investment goods) refer to all human made aids to increase production, such as technology, machines, buildings and infrastructure. Technology may be separated from capital and recorded as a separate category. Stocks, bonds and money in the bank are not capital or indeed any other economic resource. Funds derived from bank balances may be
used to acquire resources but are not themselves resources. It is quite possible for a hospital to have funding to perform additional surgery but not have the required staff or facilities to do this.

**Basic economic questions of scarcity, choice and rationing**

In determining how best to allocate resources, society must seek answers to three basic questions.

(1) *What* mix of goods and services should be produced and how much of each service should be delivered?

(2) *How* should the chosen goods and services be produced? For example, should the delivery of health services be community or hospital based? (The question of how involves efficiency and equity considerations which are discussed in Chapter 5.)

(3) *Who* should get these goods and services produced, that is, how should the wealth and income of society be shared between its citizens? (This will be explored in depth when the concept of equity is discussed in Chapter 5.) How these questions are solved depends upon the type of economy and the values and preferences of its citizens.

Because resources are scarce in relation to a society’s needs, wants and desires, society must decide how to use their scarce resources to achieve the greatest benefit for all. Some form of rationing will be necessary either by price as in a market economy, administrative means in a command economy or by a mix of both price and administrative methods in a mixed market economy.

Economies may be of three types; command or centrally planned, pure price (or market) and mixed. The two opposite ends of the spectrum are pure market economies and planned economies. However, practically all economies are mixed, falling somewhere between these two extremes. In market economies the price system dominates while in centrally planned economies choices will be made by the government. In a planned economy, most productive resources are owned by the state and allocation of resources is determined by the state. While in market economies most productive resources are privately owned and resources are allocated by the price system.

**Adam Smith: the invisible hand and the free market**

Modern economic theory owes much to Adam Smith (Smith, 1991, reprint of the 1776 book) who described the working of the market and the determination of social welfare. Smith stated that if every individual pursued their own selfish interests they would be led as if by an “invisible hand” to achieve the greatest good for all of society. "Every individual necessarily labours to render the annual revenue of the society as great as he can. He
generally indeed neither intends to promote the public interest, nor knows how much he is promoting it. He intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it. I have never known much good done by those who affected to trade for the public good."

“Through the market, greed is harnessed to serve social purposes in an impersonal and seemingly automatic way. A competitive market transmits signals to producers that reflect the values of consumers. If the manufacture and distribution of a product is profitable, the benefits it provides to buyers necessarily exceed the costs of production. And these costs in turn measure the value of the outputs that are sacrificed by using labour and capital to make the new product. Thus profitability channels resources into more productive uses and pulls away from less productive ones. The producer has the incentive to make what consumers want and to make it in the least costly way.”(Okun, 1975, p. 50)

If the price or market system is functioning efficiently, the maximum quantity of wanted goods and services will be produced at least cost from available economic resources. (Efficiency will be discussed later in this chapter). Firms will produce what consumers want, and resources will be allocated efficiently. However, free markets may fail to provide the best outcome for society and will be covered in Chapter 5: The role of Government in a Mixed Market Economy.

**The production possibilities frontier (PPF)**

The production possibilities frontier (PPF) is also called the production possibilities or transformation curve (Figure 1). It is a graph representing the maximum attainable combinations of two goods that can be produced by a firm or an economy from available resources and technology. The PPF may be used to illustrate scarcity, choice, opportunity cost and technical or production efficiency. All points on the PPF assume that resources are used efficiently, there are no unemployed resources, technology and quantities of resources are fixed and two goods or services are being compared. If any of these assumptions change the PPF will shift. Points inside the PPF are inefficient and points outside the frontier are unattainable unless technology or the quantity of resources change.
In Figure 1 opportunity cost may be illustrated in a movement along the PPF from point A to point B. In this case there would be a gain of 50 units (80-30) of health services at a cost of 31 units (91-60) of education services. Similarly, we can take any two points on the frontier and evaluate the opportunity cost of moving from one mix of outputs of education and health services to another mix. If resources that produce education services are increased maintaining all other factors constant, then the production possibility frontier will move outwards in the direction of education services. Increasing resources that produce health services will cause the frontier to move towards health services while education services would remain unchanged. If there was an increase in resources equally suitable for producing education and health, then the production possibility frontier would expand outwards to indicate a greater output of both education services and health services. Changes in technology can be explained in an analogous fashion.

The production possibilities curve is normally bowed outwards (concave to the origin) because of increasing opportunity cost to produce an additional unit of one good in terms of units foregone of the other good (Figure 2). For example, increasing units of health services will cost more in terms of units of education foregone. Similarly, there is increasing opportunity cost to produce additional units of education services with respect to units of...
health services foregone.

![Figure 2. A production possibilities curve illustrating increasing opportunity cost](image)

![Figure 3. A production possibilities curve illustrating constant opportunity cost](image)

Some resources are equally suited to producing either health or education services but other resources are best suited to producing one specific output. As production is switched from education to health, resources that are better suited to producing health services and less suited to the production of education services will be re-allocated first. Transferring production from health to education may be explained in an analogous fashion.
In practice there may be only sufficient robust information to construct a straight line production possibilities curve. For example, information may only be available for producing hundred units of education services and 0 units of health services, or 100 units of health services and 0 units of education services (Figure 3). In this situation we could assume constant opportunity costs over all levels of output, that is, a straight line PPF.

The law of diminishing returns

As increasing quantities of a variable input are added to fixed amounts of all other inputs (nurses, equipment and consumables) the additions to output gained as a result will at some point begin to decline. This might apply if a hospital attempted to increase the number of surgical procedures by increasing only the number of surgeons holding all other inputs (such as numbers of nurses and operating theatres) constant.

Economies of scale

Economies of scale occur when an increase in the scale of production results in a fall in long-run average total costs. This is because in the long run all inputs may be changed. If a number of small hospitals were merged to create a larger unit some costs (for example, administration services) could be spread over a greater output. This is a long run concept and should not be confused with the law of diminishing returns which relates to the short run where some inputs cannot be changed.

Efficiency (technical or productive and allocative)

Output is technically or productively efficient if the maximum output is attained from given resources and technology. This implies that there are no unemployed or under employed resources. If production is efficient it is possible to obtain more of one good only by sacrificing some of another. All points on the PPF are technically efficient, but not all are allocatively efficient. Resources are allocated efficiently to different activities if the maximum quantity of wanted goods and services are produced from existing resources and technology. Allocative efficiency also means that societies aggregate utility or satisfaction is maximised, but it tells us nothing about equity.

Equity

Equity refers to the allocation of society’s wealth and output between its citizens and should not be confused with efficiency. Equity and efficiency are discussed in greater detail in Chapter 5: The role of Government in a Mixed Market Economy
The nature of health and the determinants of health status

Health is not easy to define. The definition of health by the World Health Organization (WHO) is as follows: "Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity. The enjoyment of the highest attainable standard of health is one of the fundamental rights of every human being without distinction of race, religion, political belief, economic or social condition" (WHO, 1948). This definition is of some value for setting idealised targets but using it to evaluate health status is not realistic as most individuals would probably be assessed as unhealthy on the basis of the WHO criteria.

“In 1977, the World Health Assembly decided that the major social goal of governments and WHO should be the attainment by all people of the world by the year 2000 of a level of health that would permit them to lead a socially and economically productive life”.

Figure 4. Some health related quality of life dimensions

Part of the reason why a definition of health is so elusive is the multidimensional nature of health and health status. Health status may be defined in health related quality of life measures that cover a number of dimensions such as, mobility, self-care, usual activity, pain/discomfort, anxiety and depression (Figure 4). See also (Marmot & Wilkinson, 2006) and (EuroQol Group, 2013).

Determinants of health status for a healthy population include the environment, genetics, demographics, health promotion, diseases, health services, food, income, values and beliefs, education and lifestyle (Figure 5). These determinants not only influence health related quality of life status but are themselves influenced by health related quality of life.
Other authors and organisations have described similar models that focused on multiple determinants of health (Evans & Stoddard, 1990) (WHO, 2015) (Harvey & Taylor), McDowell (McDowell). The following comment from the WHO organisation is also very pertinent to this discussion. “Many factors combine together to affect the health of individuals and communities. Whether people are healthy or not, is determined by their circumstances and environment. To a large extent, factors such as where we live, the state of our environment, genetics, our income and education level, and our relationships with friends and family all have considerable impacts on health, whereas the more commonly considered factors such as access and use of health care services often have less of an impact.” (WHO, 2015).

![Diagram of determinants of health status]

**Questions to consider**

- Why is a budget increase for health not the same as an increase in health resources?
- Why are scarcity and choice and the need for rationing as much a problem for the USA as for Tonga or the Cook Islands?
- Could you use the production possibilities curve to demonstrate opportunity cost and choice within the health sector? If so how?
- Does the rate of unemployment in a region have an impact on the need and utilisation of health services? Why?
- How would you use the events that followed from Cyclone Pam which devastated much of the South Pacific in mid-March 2015, and hurricane Katrina that struck New Orleans...
on 29 August 2005 demonstrate to the impact of determinants other than health care on health status?
Chapter 2: The Market for Health Care

Objectives and Key Concepts

Important aspects of the market in relation to health are introduced and on completion of this chapter you should:

- Be able to explain the determinants of demand and supply.
- Understand the functioning of the price or market system.
- Appreciate the role of prices in coordinating production and consumption and in allocating and rationing scarce resources.
- Be able to apply your knowledge of consumer behaviour and the concepts of demand and supply to the allocation of resources in health care.
- Understand the concept gaining utility or satisfaction from consuming a good or service.
- Be aware of the law demand and a consumer’s desire, willingness and ability to pay for a good or service at particular prices.
- Comprehend the law of supply and the quantities of a good or service that supplier will be willing and able to supply at particular prices.
- Know about the interaction of supply and demand to establish the market equilibrium, price and quantity.
- Understand how the value of the coefficient of the own price elasticity of demand determines the responsive the quantity demanded is to a price change.
- Be able to explain the relationship between unmet need and excess demand.

Demand

Utility and marginal utility

Consumers demand goods because they provide utility or satisfaction and meet a perceived need or want. (A perceived need is not the same as an actual need – to find out more - see Chapter 3: Elasticities of Demand and Special Features of the Market for Health Services). Utility is the satisfaction an individual derives from consuming a good while marginal utility is the increase in (or additional) satisfaction an individual derives from consuming one more unit of a good. As more of a good is consumed, the additional utility derived from successive units of the good eventually declines – this is called diminishing marginal utility.

The law of demand

The law of demand states that, the quantity of a good demanded by buyers in a market tends to decrease as the price of the good increases (and vice versa) all other factors relating to demand held constant.
Demand is a maximum concept in that it measures the maximum price consumers are willing to pay for various quantities (or the maximum quantity consumers will demand at various prices). As the demand curve measures consumers’ willingness to pay, it also measures benefits to consumers. The demand side of the market is concerned with benefits and consumers.

**Reasons why the quantity demanded reduces as a good’s own price increases**

As the price rises, fewer consumers are prepared to purchase the good or service, and those who are willing and able to purchase, will want to buy less (Table 1). This may be explained three ways. *Income effect:* The purchasing power of the consumers’ income falls as price rises, and for a normal good this means less will be demanded. *Substitution effect:* If the price of a good rises, consumers will substitute cheaper equivalents where possible. *Declining marginal utility:* Because the first unit of a good consumed has the highest utility, and the last unit consumed has the lowest utility, consumers are willing to pay a higher price for the first unit consumed than the last unit consumed.

**Describing demand**

Demand may be described in three ways, as a schedule or table, as a curve a graph, or as a function or equation. We will use the market for physiotherapy consultations to illustrate demand, supply and market equilibrium.

**Table 1. Demand schedule**

<table>
<thead>
<tr>
<th>Price per consultation $/</th>
<th>Quantity of consultations demanded per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>3</td>
</tr>
<tr>
<td>66</td>
<td>4</td>
</tr>
<tr>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>54</td>
<td>6</td>
</tr>
<tr>
<td>48</td>
<td>7</td>
</tr>
<tr>
<td>42</td>
<td>8</td>
</tr>
<tr>
<td>36</td>
<td>9</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
</tr>
</tbody>
</table>

**Determinants of demand**

The determinants of demand may be divided into two groups; (1) changes in the price of the good in question (own price), and (2) a set of other determinates, which include; prices of related goods, consumer incomes and wealth, tastes and preferences and expectations.

*Prices of related goods:* Substitutes are goods that satisfy similar needs, e.g. jam and honey, or osteopath and physiotherapy services. Complements are goods that are consumed together, such as, beer and chips, needles and syringes physiotherapy consultations and x-
rays. An increase in the price of a substitute increases demand for the good in question while an increase in a complement reduces demand.

An increase in household incomes will cause demand to increase, while a decrease household disposable income (spending power) will cause demand to decrease. All other things equal, an increase in household wealth will cause demand to increase, and a decrease in income is likely to cause demand to decrease.

If tastes and preferences change in favour of the good in question demand will increase, and vice versa. For example, there are two forms of treatment, hospital based physiotherapy services versus community physiotherapy services. If preferences shift towards community treatment, then demand for community treatment would increase and the demand for hospital treatment would decline (provided all other determinants remained constant).

Expectations of future income, wealth and prices are important in determining consumer behaviour. If consumers expect their incomes to rise in the near future, they may desire to spend more now thus increasing demand. While expectations that future income and wealth will fall will reduce demand.

The number of potential consumers defines market size. An increase in market size will cause demand to increase, while a decrease in the potential number of consumers will result in a decrease in demand. (The determinants of demand are summarised in Equation 1).

\[ Q_d = f(P_o, P_s, P_c, P_s, I, T, W, E, N) \]

Where:

- \( P_o \) = Own price (physiotherapist fee)
- Determinants other than the price of the good in question (exogenous determinants)
- \( P_s \) = Price of substitutes
- \( P_c \) = Price of complements
- \( I \) = Household income
- \( T \) = Tastes and preferences
- \( W \) = Household wealth
- \( E \) = Future expectations
- \( N \) = Number of potential consumers

Equation 1. Determinants of demand

The difference between a shift in demand and a change in quantity demanded

Movements along a demand curve, also called changes in quantity demanded result from a change in the price of the good in question (the price of a physiotherapy consultation), all other determinants of demand held constant. Shifts of the demand curve are also called changes in demand meaning the whole demand curve or schedule shifts. Shifts in demand result from changes in a determinant of demand other than the price of the good in question. (See Figure 6).
Demand, benefits and willingness to pay

Demand measures consumers’ willingness to pay for the perceived benefits of a product or service. More specifically, the demand curve defines the maximum consumers are willing to pay for successive units of a good. The maximum that a consumer is willing to pay is also called the reservation price. The demand curve also reveals preferences and defines marginal benefits.

Supply

The supply side of the market is concerned with sellers and producers. Supply relates to costs and production. Many of the concepts discussed under demand are also relevant to supply. For example, supply may be represented in a table, by a graph and as a mathematical function.

Supply, sellers and cost

Supply is a minimum concept, it describes the minimum prices suppliers are prepared to accept for various quantities supplied.

Law of supply

The quantity of a good supplied by sellers in a market tends to increase as the price of the good increases, all other factors relating to supply held constant. As price rises new sellers enter the market and existing sellers will be prepared to sell more (Table 2).
Table 2. Supply schedule

<table>
<thead>
<tr>
<th>Price per consultation ($)</th>
<th>Quantity of consultations Supplied per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>10</td>
</tr>
<tr>
<td>66</td>
<td>9</td>
</tr>
<tr>
<td>60</td>
<td>8</td>
</tr>
<tr>
<td>54</td>
<td>7</td>
</tr>
<tr>
<td>48</td>
<td>6</td>
</tr>
<tr>
<td>42</td>
<td>5</td>
</tr>
<tr>
<td>36</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
</tr>
</tbody>
</table>

Determinants of supply may be split into two groups; the price of the good in question and a range of exogenous factors influencing supply (Equation 2).

\[ Q_s = f(P_o, P_s, P_c, I, T, W, E, N) \]

Where:
- \( P_o \) = own price (physiotherapist fee)
- \( P_s \) = Price of substitutes in production (other goods that could be made using the same resources or inputs)
- \( P_c \) = Price of complements in production (goods that are produced jointly (or as by-products) using the same resources or inputs)
- \( P_i \) = Price of resources or inputs
- \( T \) = Technology
- \( E \) = Future expectations
- \( N \) = Number of suppliers/producers

Equation 2. Determinants of supply

The most important determinant is price of the good in question (in the example being used), physiotherapist fees. All other things being equal, price changes (of the good in question) result in movements up or down the supply curve. Changes in determinants other than the physiotherapist fee cause shifts in the supply. Any change that increases productive capacity or makes production and sales more profitable induces an increase or rightward shift in the supply curve.

Prices of other goods that could be produced from the same resources. These goods are either substitutes or complements in production. Substitutes (for example, an office consultation or a home visit) and complements (for example, some consultations may generate referrals to a fitness clinic). If the price of a substitute increases, supply of the good in question will decrease because some production will be switched to the substitute. An
increase in the price of a compliment increases supply (all other things equal) because increased revenue is derived from the complement.

*The price of inputs* to the production process, for example, an increase in wage rates or a rise in the price of medicines, will, all other things being equal, reduce supply.

*Expectations* relating to the business environment. Expectations about future prices and market conditions, such as, an increase in practice subsidies or health insurance payment rates would be expected, all other things equal, to increase supply. Conversely, expectations of cost increases would decrease supply.

The *number of producers or suppliers*. An increase in the number of producers of health care services such as more physiotherapists from overseas would increase the supply of physiotherapists.

**The interaction of market forces and price rationing**

Thus far, demand and supply have been considered as separate and independent entities. Consider now, the interaction of the forces of supply and demand to determine market prices and output. Consumers always want to buy at the lowest possible price and maximise their utility. Sellers want to attain the highest prices and maximise their profits. These two conflicting objectives are resolved by interaction of supply and demand in the market place. Market equilibrium occurs where the forces of supply and demand are equalised as illustrated in Table 3 and Figure 7.

Market equilibrium may be defined as, a situation where the demand curve cuts the supply curve, the quantity demanded equals the quantity supplied, and there is no excess demand or supply. The equilibrium price is also called the market clearing price because at this price the market is cleared of excess demand and excess supply. Market prices coordinate production and consumption and ration scarce resources.
Table 3: The demand schedule, supply schedule and market equilibrium for physiotherapy consultations.

<table>
<thead>
<tr>
<th>Price per consultation $</th>
<th>Quantity of consultations demanded per day</th>
<th>Quantity of consultations supplied per day</th>
<th>Excess demand caused an upward force on price and quantity demanded</th>
<th>Excess supply caused a downward force on price and quantity demanded</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>3</td>
<td>10</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>4</td>
<td>9</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>5</td>
<td>8</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>6.5</td>
<td>6.5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Market equilibrium

| 48 | 7 | 6 | 1 |
| 42 | 8 | 5 | 3 |
| 36 | 9 | 4 | 5 |
| 30 | 10| 3 | 7 |

Figure 7. The market for physiotherapy consultations

A shift in demand or supply will result in either excess demand or excess supply and cause the equilibrium price to change. Excess demand produces an increase in quantity demanded and equilibrium price while excess supply initiates a similar but opposite chain of
events. If both demand and supply simultaneously change the impact on equilibrium price and quantity will be indeterminate.

If demand increases and supply remains unchanged (the demand curve shifts right or outwards). Excess demand will develop at the old price and there will be an unplanned fall in stock levels. This unplanned fall in stocks sends a message to sellers that they could raise prices. Sellers increase prices. Quantity demanded decreases and at the same time quantity supplied increases and price rises. The end result is that a new equilibrium with higher price and greater quantity is attained. The opposite chain of events is initiated by a decrease in demand with supply constant. Similarly, if supply changes and demand remains constant, excess demand or excess supply would result and alter the market equilibrium price and quantities.

Table 4 describes how shifts in supply and demand will change the equilibrium price and quantity of a good or service.

<table>
<thead>
<tr>
<th>Resultant changes in equilibrium price and quantity demanded/supplied</th>
<th>Price</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand increase no change in supply</td>
<td>Increase</td>
<td>Increase</td>
</tr>
<tr>
<td>Demand decrease no change in supply</td>
<td>Decrease</td>
<td>Decrease</td>
</tr>
<tr>
<td>Supply increase no change in demand</td>
<td>Decrease</td>
<td>Increase</td>
</tr>
<tr>
<td>Supply decrease no change in demand</td>
<td>Decrease</td>
<td>Decrease</td>
</tr>
<tr>
<td>Supply and demand both increase</td>
<td>Indeterminate</td>
<td>Increase</td>
</tr>
<tr>
<td>Supply and demand both decrease</td>
<td>Indeterminate</td>
<td>Decrease</td>
</tr>
<tr>
<td>Demand increase supply decrease</td>
<td>Increase</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Demand decrease supply increase</td>
<td>Decrease</td>
<td>Indeterminate</td>
</tr>
</tbody>
</table>

**Questions to consider**

- Considering the New Zealand market for nursing services, list the determinants of demand and supply.
- Explain in economic terms the consequences of introducing fees controls on specialist medical services.
- Explain the consequences of the government selling public hospitals and introducing a free market user pays system for hospital service in New Zealand.
- Why do New Zealand Customs monitor the street price of heroin? (Campion, 1999).
• Discuss three ways in which the problems of (a) increasing waiting lists for elective surgery and (b) shortages of nurses and physiotherapists in public hospitals could be solved.

• Explain why there are waiting lists for many state funded health services New Zealand and the United Kingdom?
Chapter 3: Elasticities of Demand and Special Features of the Market for Health Services

Objectives and key concepts

After studying this chapter, you should be able to:

- Understand and interpret own price, cross price and income elasticities of demand.
- Discriminate between elastic and inelastic demand.
- Measure and interpret the values of the coefficients of price elasticity of demand.
- Understand the relevance of price elasticity of demand to health policy.
- Recognise the special features of the market for health services.
- Appreciate the difference between needs and wants.

Elasticities of demand

Elasticities of demand measure the reaction of quantity demanded produced by (1) a change in the good’s own price (own price elasticity of demand), (2) an alteration in prices of related goods (cross price elasticity of demand) and (3) variations in income (income elasticity of demand).

Own price elasticity of demand (PED)

The own price elasticity of demand is the responsiveness of the quantity demanded to a change in the price of the good, holding all other factors constant. For example, we may wish to know the percentage change in the number of general practitioner consultations if the consultation fee is raised by a specific percentage or the impact on wine consumption of raising the tax on wine. The coefficient of the price elasticity of demand (PED) is the percentage change in quantity demanded of the good in question divided by the percentage change in the good’s own price (Equation 3).

\[
\text{Own price elasticity of demand (PED)} = \frac{\% \Delta \text{ quantity demanded}}{\% \Delta \text{ in own price}} = \frac{\% \Delta Q_d}{\% \Delta P}
\]

Where \( \Delta Q_d = Q_d2 - Q_d1 \) and \( \Delta P = P2 - P1 \) and \( \Delta = \text{change} \)

Equation 3. Coefficient of own price elasticity of demand

The PED almost always has a negative sign because when the change in price is positive (increased) the change in quantity demanded will be negative (a reduction). The converse is true when price falls. However, it is often convenient to ignore the price and only
consider the numerical value in determining whether demand is elastic, inelastic or unit elastic demand. If the value of the coefficient of the price elasticity of demand is known, it is possible to calculate the percentages change in quantity demanded in response to a given percentage change in price. The percentage change in quantity demanded equals the numerical value of the price elasticity of demand coefficient multiplied by the percentage change in price \( \% \Delta Q = PED \times \% \Delta P \).

If the price elasticity of demand is -5.0 a 1% increase in price will result in a 5% decrease in quantity demanded, that is: \( % \Delta Qd = -5 \times 1 = -5\% \). In another example the own price elasticity of demand for beef is -0.7 (United States Department of Agriculture Economic Research Service, 2012) which means that a reduction in price of 1% causes an increase in quantity demanded of 0.7%.

The impact of a price change on quantity demanded and revenue/expenditure is dependent on the numerical value of the coefficient. The larger the coefficient the greater the elasticity of demand and the greater the impact on quantity demanded (Table 5). If demand is elastic and price falls the quantity response is greater than the price response and revenue of sellers and expenditure of consumers increases.

<table>
<thead>
<tr>
<th>Type of elasticity</th>
<th>Magnitude of the value of the PED coefficient (ignoring the sign)</th>
<th>% change in quantity compared with the % change in price</th>
<th>Change in expenditure/revenue in response to a price change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic</td>
<td>&gt; 1</td>
<td>%(\Delta Q &gt; %\Delta P)</td>
<td>Price and revenue change in opposite directions</td>
</tr>
<tr>
<td>Unit elastic</td>
<td>= 1</td>
<td>%(\Delta Q = %\Delta P)</td>
<td>No change in expenditure/revenue</td>
</tr>
<tr>
<td>Inelastic</td>
<td>&lt; 1</td>
<td>%(\Delta Q &lt; %\Delta P)</td>
<td>Price and revenue change in the same direction</td>
</tr>
</tbody>
</table>

Where > means greater than, < is less than and \( \Delta \) is change.

**Determinants of own price elasticity of demand**

The magnitude of the coefficient of the price elasticity of demand is dependent upon a range of factors; the number of substitutes, the proportion of the consumers’ disposable income spent on the good, time and whether a good is a necessity or a luxury.

**Substitutes:** The greater the number of close substitutes the higher the price elasticity of demand because a consumer has more alternatives from which to choose.
Disposable income: The higher the proportion of a consumer's income spent on a good and the higher the price the greater the elasticity because price changes are more likely to be noticed and action taken.

Time: Demand is often more elastic in the long-run than in the short-run because over time new substitutes develop.

Necessities (and addictive substances) and luxuries: Necessities tend to have a lower elasticity than do luxuries (items that may be desired but are not actually needed). For example, basic health care would be regarded as a necessity by many people whereas high performance motor vehicles would be regarded as a luxury.

Attempting to reduce the consumption of some foods deemed to be unhealthy by imposing added taxes is frequently ineffective as most foods have a low price elasticity of demand. The magnitude of the tax increase would need to be considerable to produce any meaningful change in quantity demanded. Although taxes may not greatly influence food consumption they do enable revenue to be collected and this may be used to fund other health initiatives. Even a small tax could raise considerable revenue (Nestle & Jacobson, 2000) (W G Scott & Scott, 2014).

The authors of a review article (Green et al., 2013) provide a good example of how own price elasticities of demand differ for different foods. “… confectionery tends to have larger elasticities, as for most people it is not a necessity and also has a relatively high price, thus requiring a larger proportion of the available budget. Dietary staples, such as cereals, tend to have smaller elasticities, because these foods are necessities in the diet, are usually cheaper, and people conserve their income for spending on such essentials when prices increase. In a similar way, low income countries tend to have higher price elasticities for all foods than high income countries, because food represents a large share of total income in these countries, hence price changes have a larger impact on budget allocation.”

Income elasticity of demand (YED)

The income elasticity of demand measures the proportionate response of quantity demand (Qd) resulting from a change in real income (Y). The formula for calculating YED = %∆Qd ÷ %∆Y. The larger the value of the income elasticity of demand the greater the sensitivity of quantity demanded to a given percentage income change. For example, if the YED for beef is 0.95 (United States Department of Agriculture Economic Research Service, 2012) then a 1% increase in income causes a 0.95% increase in quantity demanded. A YED of 1.6 mean that a 1% increase in income would result in a 1.6% change in quantity demanded.
Cross price elasticity of demand (XED)

The cross price elasticity of demand relates the percentage change in the quantity demanded of one product in response to a percentage price change of another product. For example, USDA (United States Department of Agriculture Economic Research Service, 2012) data indicate that the cross price elasticity of demand for beef with respect to pork is +0.218 and bread -0.131, which means that if the price of beef increases by 1% the quantity demanded of pork increases by 0.218% and bread falls by 0.131%. If the cross price elasticity of demand is positive then the two goods are substitutes, but if the cross price elasticity of demand is negative the two goods are complements. The example quoted in this paragraph indicates that beef and pork are substitutes and beef and bread are complements.

Special features of the demand for and supply of health and health services

The Grossman model

Grossman described inputs into a production function for an individual’s stock of health with an output of healthy time (Grossman, 1972, 1992) (Figure 8). Production of health was dependent upon many environmental variables (including health care, education, housing, income, employment status, and nutrition), the most important of which was education. It was assumed that higher levels of education would improve the efficiency of health production. A good historical perspective on the Grossman model is provided by Leibowitz (2004).

![Figure 8. The Grossman model](image)

Grossman regarded health as both a capital good (good health allows workers to be more productive) and a consumption good (people desire good health). The demand for health care is derived from the demand for good health – the “rational” consumer wants good health – they do not want to consume health services. “… what consumers demand when they...
purchase medical services are not those services per se but, rather ‘good health’ ” (Grossman, 1972 p 224).

**Wants and needs, demand and supply**

Wants and needs for health care are different concepts. The need for health is wide ranging and may embrace illnesses for which there is currently no treatment. Perceived needs of consumers may be quite different from assessed need made by health care professionals and policy makers. Unmet health care need is the disparity between health services considered essential to treat or prevent an illness and the actual services received. Supply is the quantity of health services or health care actually provided.

In a ground-breaking paper Bradshaw (1972; 2005, p. 46) defined four types of need relevant to healthcare; (1) normative need, (2) felt need, (3) expressed need and (4) comparative need. (1) **Normative need** is the minimum standard of health care assessed by experts, (2) **felt need** reflects individual’s views of their own wants, and (3) **expressed need** is felt need converted into demand for health care. (4) **Comparative need** is concerned with equity such as the health care provided to individuals with similar abilities to benefit. These categories of need are not mutually exclusive or all-embracing. Additional categories are possible and there may be some overlap between classifications (Figure 9).

![Figure 9. Bradshaw’s categories of need for health care](image)

The demand for health care is derived from the desire for improved quality of life and additional time for recreation and work. Although good health cannot be traded in the market place the desire for good health creates a demand for health services for which there is a market. The demand for health services in turn generates a demand for health care resources such as nurses and medicines. This process is depicted in Figure 10.
Figure 10. The demand for health care

Questions to consider

- Would you expect the demand for prescription medicines to be elastic, inelastic or unit elastic? Why?
- Explain what you understand by “derived demand” and provide a specific health sector example?
- What is the difference between need and demand?
- Define Bradshaw’s categories of need for health care.
- Why is it important to know how to interpret the own price elasticity of demand, the cross price elasticity of demand and the income elasticity of demand?
- What are the revenue implications for producers and the effect on consumer’s expenditure if the price of healthcare services is reduced or increased?
- How could the change in expenditure and revenue generated by a price change be quantified?
Chapter 4: Competition in the Health Care Market

This chapter considers market competition and market structures that are of relevance to the delivery of health services.

Objectives and key concepts

At the end of this topic you should:

- Understand the concepts of competition and market power.
- Appreciate and be able to distinguish between different market structures.
- Realize how entry barriers allow monopolies to exist and to make excess profits.
- Be aware of the importance of competition in minimising unit costs and prices.

Market competition, market power and market capture

Market power is the ability of a firm to influence market price and output and capture the market for their own benefit. The greater the degree of competition in a market the lower the market power of buyers and sellers and the greater the chance that suppliers will be efficient, and that prices and excess profits be minimised.

Market structures

Two extreme market structures may be identified. Perfect competition, is the market state with the greatest competition and monopoly/monopsony the least. Most real world markets fall in between these two extremes and display varying degrees of imperfect competition and market power. The greater the competition in a market (all other factors constant), the lower the level of unit costs and prices.

Perfect competition

In this market state there are many buyers and sellers of identical products and an individual firm has no market power and cannot influence market price or the quantity traded. In other words, sellers and buyers are price takers. All sellers and buyers are assumed to have perfect market knowledge. Price and output levels are determined by market supply and demand. Prices are equal to unit costs and there are no excess profits.

Imperfect competition

Imperfect competitors are able to maintain prices above unit costs, and excess profits may be earned.
**Monopsony**

Market power enables a single buyer to influence prices. Many funding agencies are able to exert considerable monopsony buying market power over suppliers. All large hospitals are able to exercise some monopsony market power when negotiating prices with suppliers of goods and services such as catering and laundry. In New Zealand, the Pharmaceutical Management Agency Limited (PHARMAC) applies monopsony power when negotiating prices and contractual arrangements with pharmaceutical suppliers and community pharmacists. PHARMAC was setup to counter the monopoly power of sellers of pharmaceuticals. The Agency’s market power could be curtailed by eliminating or modifying their exemptions from the anti-competitive legislation.

**Monopoly**

In this type of market there is one seller and many buyers. Monopoly power enables a large seller to control prices. A monopolist sells goods or services that currently have no close substitutes (goods that meet similar needs and wants). The seller is able to charge different prices in different markets and/or to different consumers. Entry barriers restrict competition and are the reasons why monopolies exist. Barriers include patients, trademarks, licences and economies of scale. For example, a large hospital may be a price maker and attempt to maintain treatment prices above marginal costs. The subsequent lack of competition reduces incentives to minimise costs. This situation could be addressed by government regulation to prevent predatory pricing and ensure contestable markets with low entry barriers.

**Oligopoly**

Sellers have some, but not absolute, market power. The international pharmaceutical industry is an example of oligopoly where for each therapeutic category there is normally a small number of suppliers.

**Monopolistic competition**

In this type of market there are many competitors, each selling a slightly different (or differentiated) product, for example, branded merchandise. Competition between private sector health care providers such as surgeons, dermatologists, radiologists or general practitioners provide examples of monopolistic competition.

**Questions to consider**

- Would it be a good idea to expose the market for dental health services to competition by removing registration requirements?
• How could better information be provided to consumers? Would it be possible to provide perfect health care information to consumers?

• Would lower cost prescription medicines be provided to New Zealand citizens if patent protection for pharmaceuticals was removed?

• Is it possible to achieve prefect competition in the market for healthcare?
Chapter 5: The role of Government in a Mixed Market Economy

This chapter introduces some fundamental concepts in welfare economics and public policy. The focus is on the interaction between government and market forces in providing and funding healthcare services in a mixed market economy.

Objectives and key concepts

There is a broad range of public policy objectives embracing a range of economic, social, or political dimensions. Public policy goals include:

- Improving allocative efficiency.
- Raising the rate of economic growth.
- Increasing social welfare.
- Achieving an equitable distribution of income.
- Maintaining economic stability.
- Achieving a low (but above zero) inflation rate.
- Reducing unemployment.
- Improving the nation’s health (is but one of these competing objectives).

At the end of this chapter you should:

- Be acquainted with the goals and instruments of public policy.
- Appreciate the need for tradeoffs between goals.
- Be aware of the conditions necessary to achieve allocative efficiency.
- Understand why a state of allocative efficiency is desirable.
- Appreciate the concepts of equity and income distribution.
- Recognise the special nature of public and merit goods.
- Appreciate the special nature of the health care market and the need for government intervention to correct for market failure.
- Understand the concepts of social welfare, allocative efficiency.
- Appreciate the Kaldor-Hicks criterion in judging when allocative efficiency is attained.
- Be familiar with the equity/efficiency tradeoffs that may exist in the provision of healthcare services.
Government intervention in the economy

Economic justification for government intervention in a mixed market economy is to correct for (or prevent) market failure, to address equity and social justice, or to smooth economic fluctuations. Governments may impose regulations, taxes, introduce a national health insurance scheme, subsidise private provision of health care, or establish a national health service.

Maximising total social welfare and allocative efficiency

Allocative efficiency is achieved if for a given income distribution, social welfare is maximised. This means that society’s resources are allocated so that the maximum quantity of the most wanted goods and services are produced from available resources and technology. Maximising total societal welfare involves maximising the sum of the welfare of all society irrespective of equity or of income distribution. This is the utilitarian concept of maximising utility and social justice (Stiglitz, 2000, p. 100) and is an implicit assumption of economic evaluation.

Production is efficient if additional production of one good cannot be achieved without giving up some output of another good. Maximum output is produced at least cost and there are no unemployed resources. Product mix is efficient if the mix of goods produced by the economy reflects consumer preferences and has the maximum value for society (aggregated social welfare or utility is maximised). Exchange is efficient if the basket of goods produced is distributed efficiently. Allocative efficiency implies that there is perfect competition in all markets all economic units act in self-interest and everything tradable is “owned” and controlled by its owner. If allocative efficiency is to be achieved; production, product mix and exchange must all be efficient.

It is difficult to judge when an economy achieves allocative efficiency. Pareto provides a solution. Pareto’s social choice rule is as follows; for any given income distribution, if allocative efficiency is achieved, it is not possible to make anyone better off (as judged by the individuals themselves) without making someone else worse off (Pareto, 1927). Allocative efficiency is also called Pareto optimality. There is often more than one Pareto optimum possible from a given set and quantity of resources and not all Pareto efficient outcomes would be considered desirable if achieving them caused too great a departure from society’s other objectives.

The compensation principal

In practice most resource allocation changes make some individuals better off and others worse off and the Pareto rule as discussed cannot be used to evaluate changes in efficiency. The potential Pareto or Kaldor-Hicks criterion enunciated separately and in different forms by both authors (Kaldor, 1939) and (Hicks, 1939) is a partial answer to this
problem and is adopted in cost benefit analysis. Under the Kaldor version of the compensation principal, if it is possible for the gainers to compensate the losers for their loss, the project is desirable. The Hicks version of the compensation principle requires that the potential losers be able to bribe the gainers to maintain the status quo. Thus, if the benefits exceed the costs of a project and if compensation is costless then the project is desirable.

**Market failure**

Markets are said to fail if the uncontrolled market fails to achieve allocative efficiency. In addition markets may not achieve solutions that meet society’s goals of, equity and stability (Stiglitz, 2000). Culyer provides specific information on market failure relating to the health sector (Culyer, 2015).

The following situations if present will cause market failure:

**Imperfect competition and market power**

A seller (monopolist supply side market power) or a buyer (monopolist demand side market power) have the power to control prices and output. This may be corrected in part or in full by government supply/purchase of health services and by regulation. PHARMAC (Pharmaceutical Management Agency) single desk pharmaceutical purchasing is an attempt to counter monopoly power of the pharmaceutical industry.

**Immobility of resources**

Resources may get locked into an inappropriate market structure. For example, it may be more efficient if a major base hospital is relocated closer to its target population. However, if adequate funding is not available to rebuild and move resources from one site to another then the restructuring will not eventuate. Government may help by providing funding and resources to assist with restructuring of health capital.

**Imperfect (asymmetric) market information and agency relationships**

Asymmetric information occurs when one party to a transaction has better information than another party. Health care providers, purchasing/funding and delivery agencies of health care usually have much better information about health and health care interventions than consumers. Most consumers have insufficient information to enable them to make informed decisions on the safety of medicines and competence of providers. Consumers frequently rely on health professionals to act as agents on their behalf and work in their patients’ best interests. However, when an agency function breaks down, one party may capture the market for their own ends. Government could also correct for this situation by acting as an agent on behalf of society and deciding which services to purchase and/or provide to citizens.
Government registration of medicines and health care providers, and the requirement to provide information about treatment options is intended to correct for asymmetric information and allow consumers to make informed decisions. For example, if health care providers have more and better quality information than consumers, and if the health service/product (medical consultation, medicines, surgery, medical products) appears to the consumer to be of higher quality and more appropriate than it actually is then consumers will demand more than they would if they had better information. In other words, the “uninformed” demand curve lies above and to the right of the “informed” demand curve. The end result is that consumers will pay a higher price and consume more of the service than with the perfect information situation (which would be allocatively efficient if all other conditions of perfect competition were met). This phenomenon has been called supplier induced demand. Situations could also occur where the “uninformed” demand curve is below and to the left of the “informed” demand curve (W Guy Scott, Scott, & Auld, 2005).

Public information campaigns, provision of a health information website are examples of providing balanced information. Regulation to force providers to disclose adequate and accurate information may also be required.

Uncertainty

Imperfect information arising from uncertainty about the future need for health care could be moderated by state provision of health insurance or a government health service. The Accident Compensation Commission (ACC, 2016) is an example of government insurance against accidents in New Zealand and the NHS in the UK is an example of a government funded health service.

Incomplete or absent markets and managed competition

Although people want good health and a good quality of life, there is no market where these can be bought. The demand for good health generates the demand for health care goods and services but these do not necessarily improve health.

Governments may also attempt to manage competition by introducing elements of the market into public health systems. Splitting the funder from the provider (Upton, 1991) and encouraging competition between public and private health providers is one such example. The UK National Health Service also operates an internal market (Light, 2001) where health authorities (the funders) seek tenders for the provision of health services from public and private health providers. Other examples of managed competition include general practice fundholders and health maintenance organisations. Enthoven discusses the concept of managed competition with respect to the US health system (Enthoven, 1993).
Moral hazard

Moral hazard occurs when the behaviour of a person changes if they are not required to meet the full cost of their actions. Individual consumers, when they buy health insurance, or are covered by a government health service or scheme frequently do not meet the full costs of the health care they access. There is a danger that these individuals may take less care of themselves, access a different type of treatment, or consume more health care than they would if they paid the full cost out of their own pockets. If patients had to pay the full cost of hospital treatment, they may choose to have day surgery rather than treatment as an inpatient. Similarly, health care professionals may suggest different treatment options in the absence of health insurance or government funding of health. Thus, moral hazard may increase demand for health services creating budgeting and rationing problems for funding agencies, insurers and providers. State funded health insurance, nonmarket rationing of health services, introduction of co-payments and part charges payable by patients are attempts by government and insurers to counter moral hazard.

Adverse selection

Adverse selection occurs when either a buyer or a seller has information that is not known to others and uses this knowledge to capture the market for their own advantage. Health insurers in an unregulated free market have an incentive to reduce their costs by selecting low risk clients and declining to cover those likely to fall ill and require health care. A good example of adverse selection is provided by a New Zealand newspaper quote from a health insurance company. “…Tower chief executive John Minto said his company did not plan to poach Southern Cross customers because many had expensive health problems that were likely to cost Tower money… If we take on board the wrong sort of business, we'll destroy value for shareholders as well as existing policyholders….” (Lacon, 2002). This situation could be avoided by government providing health insurance for all or by regulating private insurers.

Externalities

The presence of externalities (both external benefits and external costs) cause allocative inefficiency and are a justification for government intervention. These costs and benefits (although relevant to society) do not enter into the decision making process of producers, sellers and consumers. Vaccinations for infectious diseases are examples of external benefits. Externalities may be avoided or the effects moderated by public funding subsidies and taxes.
Type of goods and market failure by the uncontrolled market

Rivalry/non-rivalry

Rivalry occurs when a person’s consumption of a good reduces the supply for others. If one person buys a pack of aspirin the same pack is not available for supply to anyone else. Non-rivalry is present if a person’s consumption does not reduce the supply for others. If one group of diners benefit from eating safe food in a restaurant because of environmental health legislation and enforcement, it does not prevent others deriving full benefits from the environmental health services.

Exclusion/ non-exclusion

Exclusion or the free rider problem means that the benefits of a good cannot be enjoyed without payment while non-exclusion is the inability to exclude non-payers. These goods will not be supplied by the free market. Government intervention is required to achieve allocative efficiency.

Private goods and pure public goods

Private goods may be provided by the price system because they are subject to rivalry and non-payers can be excluded. Pure public goods are not subject to exclusion or rivalry. They are consumed collectively. They are generally regarded by society as desirable but uncontrolled private markets will fail to provide them because of the inability to extract payment. However, the state acting as an agent for society may pay for and/or supply pure public these goods. Most environmental and public health initiatives fall into this category. The provision of clean safe drinking water and the safe disposal of sewage in a city is an example of this.

Merit Goods

Society considers that people should consume an adequate quantity of merit goods regardless of ability to pay and regardless of whether they are wanted by individual consumers. As these goods have beneficial externalities, both equity and efficiency considerations justify state funding and enforced consumption of these goods for all citizens.

A specific example of a merit good is the treatment of tuberculosis. In New Zealand, treatment for this condition is free of charge and if a person will not accept treatment it is forced upon him/her. In 1995 a young man with tuberculosis was treated in a prison hospital because he would not complete a course of treatment in Wellington Hospital and Hutt Valley hospital (Dominion reporter, 1995). It was considered that the risk to society of infecting others and the development of drug resistance strains of tuberculosis was too high to let the disease go untreated.
Unemployed resources and market failure

Unemployed resources mean that a greater output of desired goods and services could have been produced for the benefit of society. For example, a hospital may not be able to fully employ all of its high technology if it does not have enough technicians to operate the equipment. Possible solutions are for government to train more technicians or to facilitate immigration of skilled workers.

Market shortcomings other than market failure

Market shortcomings other than market failure frequently relate to aspects of equity and economic stability. Lack of equity is concerned with the distribution of income and wealth and may be moderated by public funding and supply of health services. Economic instability (prices, wages, output, business cycles) may controlled by discretionary fiscal and monetary policy interventions by government.

Equity goal

Equity and equality are not identical. Equality treats all people equally regardless of their situation and needs. Using the analogy of cutting and apportioning a cake, equality would dictate that everyone would receive the same sized slice. Value judgements on how much each individual should get are not part of the distribution decision.

Equity is concerned with “fairness and justice” and value judgements are necessary in making distribution decisions. In this situation individuals may receive different sized slices of the cake depending upon their situation. Frequently populations may have equity of access to health care but not equity of utilisation. The inverse care rule which states “...the availability of good medical care tends to vary inversely with the need of the population served. ...The force that creates and maintains the inverse care law is the operation of the market...” (Hart, 1971 412). As those with greatest need are least able to access health care in a market driven health system, Hart advocated removing as much of health care as possible from the market.

The concept of equity is further complicated by the classification of equity into horizontal versus vertical equity (McDaniel & Repetti, 1993; Wagstaff, Van Doorslaer, & Paci, 1991). An additional difficulty is described by Musgrave (reported in McDaniel & Repetti, 1993, pp. 608-609) who argued that the prerequisites of horizontal and vertical equity were “but different sides of the same coin” and “that in an ideal world, arrangements that satisfy vertical equity also satisfy horizontal equity”.

*Horizontal equity* is the equal treatment of equals. Individuals with equal ability to pay are liable for equal amounts of tax and those with an equal need for health care have equal access to health services. For example, the state may fund, or provide free of charge,
certain basic or core health services (National Advisory Committee on Core Health and Disability Support Services, 1993) (Campbell, 1995) to all citizens. Universal benefits paid to those 65 years and older are another example of horizontal equity.

*Vertical equity* relates to fairness in the unequal treatment of unequals. People with the same health need are treated in the same way. Individuals with greater incomes pay more tax than those less able to pay, and those with greater ability to pay for their health care may receive a lower government subsidy. Targeted benefits are an example of vertical equity. If core health services are not to be subsidised or supplied free of charge to all citizens the selection and targeting could be based on family income. Defining core health services is a difficult and politically charged task and there is always pressure to widen the core.

**Equity/efficiency tradeoffs**

Often a policy change will result in “winners” and losers. It is possible to increase total utility of society improving utility for some citizens and reducing utility for others. Accordingly, when planning and implementing public policies it may be desirable to trade off efficiency against equity. The efficiency/equity tradeoff must be confronted irrespective of whether the economy is growing or contracting. Using our cake analogy, if the cake is growing as a result of more efficient baking technology the smallest slice may now be larger and distribution decisions less contentious. Closing a small rural hospital and delivering hospital based health services from a larger regional hospital may be more efficient for the District Health Board (Figure 11). However, in addressing social justice issues in this case it may be necessary to subsidise transport from rural communities to the regional hospital.

![Figure 11. Equity efficiency tradeoff](attachment:image.png)

A serious problem in making these tradeoffs between equity and efficiency is that each time distribution of income is changed a new allocative efficient solution emerges. This means that efficiency and equity decisions are interrelated and cannot be made independently of each other. In confronting the tradeoff between equity and efficiency it is not sufficient to first target efficiency in the economy and then address equity. If some of the additional income generated by a more efficient economy is transferred from rich to poor the process is
not costless, some income will be lost in the transfer. “The leaky bucket” analogy (Okun, 1975, p. 90) provides a good example of income lost if a leaky bucket is used to redistribute income.

**The second best and government failure**

**Second best solution**

Public health care policy interventions aiming to maximise social welfare are often concerned with attempting to correct for market failure by attempting to fulfil as many of the allocative efficiency conditions as possible. However, given the difficulties of determining when allocative efficiency is attained, and of satisfying all the conditions for allocative efficiency (first or best choice) the concept of second best becomes relevant and important. If a Pareto efficient outcome cannot be attained, is it best to satisfy as many of the Pareto conditions as possible or would the second best (Lipsey & Lancaster, 1956) option attained in some other way be better? For example, it may be better to counter monopoly power of sellers with monopsony power of consumers rather than attempt to shift the market conditions and structure towards perfect competition. It could be argued that this is the situation relating to the pharmaceutical market in New Zealand where monopoly power of the pharmaceutical industry is met by the monopsony power of the Pharmaceutical Management Agency (PHARMAC).

**Government failure**

Frequently government involvement in the market fails to improve efficiency and may reduce allocative efficiency. Intervention policies may themselves increase inefficiencies and fail to maximise social welfare. Accordingly, governments should intervene only if the benefits of intervention exceed the costs. When inefficiencies are small it may be better for government to do nothing. For example, in an attempt to increase competition in the shipping industry a government may establish and operate its own shipping line, however the lack of performance incentives for management in the government enterprise may cause greater efficiency losses than those from lack of competition.

**Questions to consider**

- Given the many determinants of health status, is it possible that additional resources allocated to health care (if it reduces resources allocated to education and social welfare) could result in lower average health status for the population? Why?
- Do you think that the consumer is always the best judge of what will maximise their utility?
- Why have health policy makers in New Zealand been cautious in embracing the market as a solution to improving efficiency and equity?
- Can we treat the market for health care in the same way as the market for hamburgers?
• Achieving greater allocative efficiency may result in greater inequality. Low income households generally have more health problems and lower health status than higher income households. Thus increasing inequality could lower average health status for the population as a whole which in turn will lower productivity and economic efficiency. What is your solution to this dilemma?

• Why is there a problem if we first maximise allocative efficiency and then redistribute income to achieve a more equitable outcome?
Chapter 6: Health Policy and International Comparisons of Health Expenditure and Gross Domestic Product

Objectives and key concepts

At the end of this chapter you should:

- Understand the formulation of health policy in a mixed market economy.
- Identify differences in funding and delivery of health services in different countries and health systems.
- Appreciate country by country differences in performance in achieving good health outcomes.

Health policy in a mixed market economy

Key determinants of health policy, shown diagrammatically in Figure 12 should all be taken into account if public policies are to be successful and objectives achieved. The schematic is simplified in that not all influences are shown. The arrows indicate the direction of causality.

Figure 12: Determinants of health policy
Objectives of the New Zealand health and disability sector

Treasury state (The Treasury, 2014) that the New Zealand health system has three interrelated objectives (Figure 13). Its three dimensions are:

- Improved quality, safety and experience of care.
- Improved health and equity for all populations.
- Best value for public health system resources.

The New Zealand Ministry of Health have established a set of more specific health targets which are annually reviewed to take account of changing priorities. The overall objective of these targets is to improve the health for New Zealanders. Six key targets were identified (New Zealand Ministry of Health, 2015).

- Shorter stays in emergency departments.
- Improved access to elective surgery.
- Shorter waits for cancer treatment.
- Increased immunisation.
- Better help for smokers to quit.
- Better diabetes and cardiovascular services.

It is not possible to simultaneously achieve all of these objectives and this inevitably leads to tradeoffs between options. In establishing the health budget government must make tradeoffs between competing activities such as education and social welfare. Once the health budget is defined, tradeoffs must then be made between various health policies.
Social policy and economic analysis should be integrated. Decision makers and agenda setters in the formulation and evaluation of policy should adopt a societal viewpoint as the principal analytic perspective for evaluating efficiency but, in addition, determine how the policy will impact on different stakeholders. Tradeoffs between equity and efficiency goals and between the interests of different stakeholders, must be confronted and policy adjusted if necessary. Finally, the effectiveness of policies in achieving equity and efficiency objectives must continue to be evaluated as part of the policy cycle.

Economic evaluation provides one set of information for policy formulation, decision making, and monitoring but it cannot set the policy agenda, make policy decisions or implement policy. Figure 14 shows the relationship that economic evaluation has within the policy cycle.

![Figure 14. The policy cycle](image)

**Funding of healthcare in alternative health systems**

There are two extremes of health systems; pure public versus pure private. In pure public systems the state funds and supplies all health services out of taxes or compulsory social insurance. In pure private systems all health care is delivered by the private sector with users paying directly out of pocket or indirectly through voluntary private health insurance. However, most systems are mixed, falling somewhere in between the two extremes.

**International comparisons of health spending**

International comparisons indicate that the level of health expenditure per capita in a particular country is highly correlated with the level of Gross Domestic Product (GDP) per capita. The correlation is evident when current health expenditure in US dollars per capita is plotted against percentage GDP per capita, and a trend line fitted to the data shows. Countries lying on the trend line are spending the level of health expenditure per capita that would be expected given those countries GDP per capita. Countries with health expenditure per capita above the trend line are spending more than would be expected given the level of GDP per capita in those nations, the converse is the case for countries lying below the trend line. New
New Zealand is situated just above the trend line. The United States spends more on health per capita than would be expected given its level of GDP per capita, while Luxembourg spends less than would be expected from its level of GDP per capita. New Zealand Ministry of Health, 2013, p. 24 Source: OECD Health Data 2011) (Figure 15). There are differences in the levels of health expenditure between the countries and between systems with respect to level of health expenditure (public and private) but there is no “right” or optimal level of spending per capita or as a % of GDP.

Figure 15. Health expenditure per capita versus GDP per capita

A WHO study of both developed and developing countries suggests that health expenditure in general does not grow faster than GDP after taking other determinants of health expenditure into account (Kea, Saksenaa, & Hollyb, 2011). However, historical data for 30 OECD more wealthy countries show health care spending increasing at a higher rate than economic growth (Huber & Orosz, 2002).
Although higher GDP per capita\(^1\) is correlated with higher life expectancy (Figure 16), many other factors are also important in determining health status. If life expectancy in years is plotted against health spending per capita (in US dollars per capita using purchasing power parity as a conversion factor) it is seen that as health expenditure rises life expectancy from birth in years improves (Figure 16). However, this is not a linear relationship, for example, a $1000 of additional health spending per capita at low levels of GDP per capita generates a greater response in the improvement in life expectancy at higher levels of GDP per capita. Two outliers, the United States and Russia are not achieving the life expectancy that the spending on health would be expected to generate. It could be argued that much of the spending on health in the United States results from excessive diagnostic testing in response to fears by health service providers of litigation. In the case of the Russian economy it possible that those in greatest need of health care are least able to access appropriate health services (particularly health care is delivered by market forces). This is the inverse care law in operation (Hart, 1971).

**Figure 16. Life expectancy versus GDP per capita and health spending per capita**

Countries above the regression line (of which New Zealand is an example) have a better life expectancy than would be expected given the level of per capita expenditure on

\(^1\) “the purchasing power parity (PPP) exchange rate - the rate at which the currency of one country would have to be converted into that of another country to buy the same amount of goods and services in each country. To understand PPP, let’s take a commonly used example, the price of a hamburger. If a hamburger is selling in London for £2 and in New York for $4, this would imply a PPP exchange rate of 1 pound to 2 U.S. dollars. This PPP exchange rate may well be different from that prevailing in financial markets (so that the actual dollar cost of a hamburger in London may be either more or less than the $4 it sells for in New York). This type of cross-country comparison is the basis for the well-known “Big Mac” index, which is published by the Economist magazine and calculates PPP exchange rates based on the McDonald’s sandwich that sells in nearly identical form in many countries around the world.” (Callen, 2012)
health (and vice versa). Figure 16 “shows that New Zealand performs well: it has relatively high life expectancy (11th among 39 countries) for comparatively modest expenditure (20th among 39 countries)” (New Zealand Ministry of Health, 2013, p. 24 Source: OECD Health Data 2011).

Funding health care in New Zealand

In 2010 83.2% of all health funding was derived from the public sector (Figure 17) (New Zealand Ministry of Health, 2012). Government funding channelled through the Ministry of Health provided 72.5% of public funding for health. Most of the Ministry of Health’s financing is delegated to the 20 District Health Boards (s). They are responsible for providing hospital-level care and for purchasing most primary and community health services. DHBs fund health care services to a geographically defined population. Each District Health Board (DHB) negotiates contracts with health service providers, (including public hospitals, not-for-profit health agencies, iwi groups and private organisations) to supply health care that meets the health needs of the DHB population. Funding is also derived from voluntary insurance and out-of-pocket payments.

Government subsidies for health services for New Zealand citizens include; general practitioner consultations prescription items and diagnostic tests. Copayments are charged on some health services to minimise moral hazard with respect to overconsumption. For example, there is a copayment of $5 per prescription item. Dental care for adults is not subsidised. Government also provides free of charge inpatient and outpatient public hospital services. To avoid waiting lists, some patients may elect to pay (via private health insurance or out of pocket) for specialist consultations and private hospital treatments. Treatments for injuries arising out of accidents are funded (in full or with a copayment) by government through the Accident Compensation Commission (ACC).

![Figure 17. % shares of New Zealand’s total current health funding, 2000 and 2010](chart.png)

Service providers include: pharmacists, laboratories, radiology clinics, primary health organisations, general practitioners, midwives, voluntary providers, community trusts, private
hospitals, Māori and Pacific providers and disability support services. General practice and aged care services are funded by DHBs but provided by private sector groups. PHARMAC negotiates prices on behalf of the DHBs for the purchase of pharmaceuticals and vaccines. ²

New Zealand government has reciprocal health agreements with Australia and the United Kingdom (UK). For New Zealanders, visiting these countries, certain services may be publicly funded to the same extent as for a national of the country they’re visiting. (New Zealand Ministry of Health, 2011)

Questions to consider

- Is one health delivery system clearly “better than another”?
- What criteria would you use to rate a health system?
- Does spending more on health care improve the health status of a nation’s population?
- Why did the health sector reform process commence in New Zealand and in other countries?
- Did the health sector reforms in New Zealand produce efficiencies (with respect to lower costs) and improved health status?
  - What happened to equity, access to, and uptake of health services?
  - Did this also occur in other countries?
- What is the relationship between per capita spending on health as a % of GDP and per capita GDP in developed countries?
- How much as a % of GDP does New Zealand spend on health? Is this too low, too high or just right? Why?

² “PHARMAC makes choices about District Health Boards’ (DHBs’) spending on vaccines, community and cancer medicines. PHARMAC also makes decisions about the medicines funded in DHB hospitals and is working towards budget management of hospital medicines and medical devices” (PHARMAC, 2016).
Chapter 7: Costs and Effects of a Health Care Intervention or Policy

Objectives and key concepts

This chapter introduces health costs and effects that are relevant to the delivery of health services. After reading this chapter you should:

- Understand opportunity cost.
- Recognise that not all costs are identified by money payments.
- Appreciate the different classifications of costs in health economics.
- Comprehend the range and difficulty of measuring both costs and health effects.
- Be able to apply the concept of incremental analysis.
- Understand the importance of perspective.

Costs and Money

*Explicit* costs involve money payments while *implicit* costs are not associated with money payments when the costs are incurred. An example of an *explicit* cost would be wage payments to employed staff. An example of an *implicit* cost is the cost to a hospital of using a self-owned building for its administration offices. No money payments are made but there is an opportunity cost because the building space could have been rented to a business. The magnitude of this cost in dollars could be estimated by reference to office space rentals in the surrounding area.

Opportunity Cost

Opportunity cost may be defined as the value of the next best alternative foregone when making a choice. The cost of choosing more hip replacements could be fewer coronary bypass operations. In other words, opportunity cost is the value of the benefits foregone in choosing one option rather than another. Opportunity cost is a concept relevant to both explicit and implicit costs.

Opportunity costs are not necessarily measured by market prices. If a particular good or service is not traded on a market, or a market for a good does not exist there is no market price. Markets may also fail to yield prices that reflect opportunity costs. For example, a health care provider relying on voluntary labour may not charge clients market rates.

Money payments may (or may not) be involved with opportunity costs. Consider a business enterprise that owns the building it uses. The business will not be paying rent (no money payment) but the use of the building has an opportunity cost because it could be rented to someone else. In another example, the opportunity cost of getting a haircut could
consist of wages foregone (a money payment not received) plus the money paid to the hairdresser. The money paid for the haircut could have been used to buy something else and the time taken to get the haircut could have been used for paid work or some other activity.

Examples of opportunity costs:

- If a nurse chooses to work weekends rather than be with her/his family, the opportunity cost of working weekends is the value of the benefits that would have been derived from being with her/his family.
- A large business enterprise that decides to use some of its revenue to build a new in-house sports and leisure facility for its staff. The opportunity cost of this decision is the value of the benefits (for example, interest foregone) it would have achieved had the next best alternative use for the money been to deposit it in a bank to earn interest.
- The New Zealand government annual budget allocates $100 million more to education rather than to health care. The opportunity cost of this policy decision is the value of the benefits that would have resulted from the additional expenditure on health care.

**Perspective and differences between decision-makers**

The opportunity cost of a choice may differ between decision-makers. Different decision-makers may have different perspectives and alternatives available to them, and a particular option may not be valued in the same way by all decision-makers. In making a choice of either going to the movies or doing something else, the next best alternative for one person may be going to a club while for another it may be watching television at home. In another case, two people may have as their next best alternative, going to a cafe, but they may each value this option differently.

![Figure 18. Perspective and costs](image-url)
There is a wide range of perspectives or interest groups relevant to the provision, funding and consuming of health services. Four such perspectives are, patients and their families, providers, funders and the all-inclusive societal perspective (Figure 18). Frequently is derived from a mix of interest groups including insurance companies, government agencies, voluntary societies and out of pocket payments by the patient.

**Internal or private costs and benefits**

Costs and benefits that are relevant to one cluster in society may not be applicable to another. Each subset of society considers only those costs and benefits (internal or private costs and benefits) that fall within its boundaries. One set of decision makers will not necessarily see or take account of the impact that their actions will have on other decision-makers. Internal or private costs and benefits relate to a particular cluster or individual and are based on alternatives available to that subset or individual. Internal costs applicable to a private hospital could include medicines and other consumables, nursing, plant and equipment insurance and interest on debt. Judgements based on private (internal) costs and benefits may not result in an outcome consistent with decisions that would be made by society using social costs and benefits.

In a private market, demand measures benefit to consumers. Consumers will continue to purchase a good or service until the price of the last unit purchased is equal to additional utility derived from that unit, that is, until price equals marginal utility (\(P = MU\)). Supply measures cost to suppliers. Firms continue to expand output until the additional revenue derived from the last unit is equal to its cost, that is, until price equals marginal cost (\(P = MC\)). Provided markets are efficient the equilibrium price equals marginal utility which equals marginal cost (\(P = MU = MC\)).

In New Zealand, most prescription medicines consumed by citizens are prescribed by general practitioners, dispensed by pharmacists, and are fully or partially funded by the government. Thus, there are four relevant perspectives: (1) consumers, (2) prescribers, (3) dispensers, pharmaceutical distributors and manufacturers, (4) funders. Each subset within society has its own agenda, its own set of internal costs and benefits, and each attempts to maximise its net benefits and shift costs to another group.

**Social (or public) costs and benefits**

Social costs and benefits relate to society and are based on alternatives available to society. Societal cost is the sum of all costs, private and external. The societal perspective is the only perspective that captures all costs and benefits of a decision. The optimal solution for society is the output level where marginal social benefit (MSB) equals marginal social cost (MSC).
Externalities or spillovers (external costs and benefits)

Externalities are costs or benefits imposed by a decision on others not involved in making that decision. If no compensation is paid or received, these externalities are not taken into account by decision makers. Beneficial externalities are also called positive externalities or external benefits, while detrimental externalities are also called negative externalities or external costs. Where external costs occur the private market supply curve will not capture all of society’s costs resulting from the activity. The result is that too many of society’s resources are devoted to the activity and costs have been transferred from one subset of society to another. Figure 19 illustrates costs generated by a chemical factory.

![Figure 19. External costs diagram](image)

Escaping external costs would be captured if the supply curve was to be raised by the amount of the externality. This new curve is called the marginal social cost curve and all
costs are now captured. One way of moving the supply curve upwards is to impose a per unit tax on those suppliers causing the external cost. See Figure 20.

Examples of external costs and cost shifting/ transfer costs

Where a medical specialist works both in a private clinic and as a consultant in a public hospital, low profit margin cases may be directed to the public system and high profit cases captured by the private clinic. Costs are thus shifted from the private clinic to the public health sector.

Consider the rather narrow perspective of a public sector provider and funder of hospital based services. From the perspective of the funder, closing all small rural hospitals and requiring all patients to travel to a large base hospital may reduce costs through economies of scale. Hospital costs would fall but external costs (travel costs and production loss) would rise. Costs have been shifted from the government funding agency to the patient and family and possibly employers if paid time off work was granted.

If a societal perspective is adopted these external costs would be taken into account in the decision making process.

External benefits

Where external benefits occur the private market demand curve will not capture all of society’s benefits generated by the activity and too few of society’s resources will be allocated to the activity. External benefits could be captured if the demand was to be raised by the amount of the externality. This new curve is called the marginal social benefits curve and all benefits are now captured. One way of moving the demand curve upwards is to introduce a per unit subsidy paid to consumers. (See Figures 21 and 22).

Figure 21. External benefits diagram
Vaccination for childhood infectious diseases provides an example of public goods and external benefits. If vaccinations were provided by the free market not all families would want to or be able to buy them for their family. Buyers of vaccinations would confer a benefit not only to their own children but also to the free riders because the risk of childhood infectious diseases spreading to unvaccinated individuals would be reduced. Under a free market system less than the socially optimal quantity of vaccinations would be supplied. The state could rectify this problem by supplying the vaccinations to all, free of charge, or by providing a consumer subsidy (funded from taxes levied on all citizens). Another example of an external benefit is herd immunity (John & Reuben, 2000) (Betsch, Böhm, & Korn, 2013) that occurs when the majority of the at-risk population is vaccinated and epidemics are avoided. For example, 95% vaccination coverage would eradicate measles in Europe.

Cost classifications in health care

There are a number of different ways in which costs may be classified. British and Canadian authors (M.F. Drummond, O'Brien, Stoddart, & Torrance, 1997; Michael F Drummond, Sculpher, Torrance, O'Brien, & Stoddart, 2005) delineate costs by sector or perspective as: health sector, patient and family, and other sectors. A US authority (Luce, Manning, Siegel, & Lipscomb, 1996) classifies costs as either direct costs (direct health care costs and direct non-health care costs and the value of patient time for treatment), or productivity costs (changes in production or output). A variant of this method divides costs as follows: avoided costs (costs generated by the condition that may be avoided by treatment), direct costs, indirect costs (United States National Library of Medicine National Information Center on Health Services Research and Health Care Technology (NICHSR),
I choose to classify costs as direct medical, direct non-medical, indirect and intangible because it enables cross tabulation of cost type by perspective (Table 6 and Figure 23).

Table 6. Cost classifications by perspective

<table>
<thead>
<tr>
<th>Cost classification</th>
<th>Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Health sector</td>
</tr>
<tr>
<td>Direct medical</td>
<td>#</td>
</tr>
<tr>
<td>Direct non-medical</td>
<td>#</td>
</tr>
<tr>
<td>Indirect</td>
<td>Σ</td>
</tr>
<tr>
<td>Intangible</td>
<td>Σ</td>
</tr>
<tr>
<td>Total cost</td>
<td>Σ</td>
</tr>
</tbody>
</table>

Note. # in a cell indicates the cost type is relevant to that particular perspective
Σ denotes a row or column total

Figure 23. A commonly used classification of health care costs

Direct costs

Direct costs are directly related to health services used to treat or prevent an illness/accident but the services may not be delivered by the health sector.

Direct medical costs of an illness/accident are the opportunity costs of resources directly utilised by treatment or prevention and are concerned with services delivered by the health sector. These costs are usually explicit (involving money payments). Although these costs fall directly within the health sector they may be funded from a range of sources.

Direct non-medical costs are directly related to treatment or prevention of illness/accident but the services may not be provided by the health sector. Some of these costs are
implicit (not identified by money payments). Such costs are generally borne by patients and their friends and families and are concerned with services provided by non-health sector. Examples include; transport (ambulance, private motor vehicle, taxis and public transport) and other additional costs as a result of the illness or intervention (special foods, clothing, wheel chairs, walking sticks, hearing aids, glasses, prostheses, housekeeping, modifications to home or motor vehicle, and hotel costs if the patient must travel away from home for treatment).

**Indirect and intangible costs**

*Indirect costs* (productivity costs) are the time-related costs arising from illness as a result of reduced capacity to work or to engage in leisure-time activities. They also include lost production to society through death. Indirect costs are external to the health sector and fall on patients and the rest of society. They are relevant to both paid and unpaid productive activity.

*Intangible costs* are the costs of pain, suffering, and the value placed on the loss of life. Although changes in quality of life and many of the costs of lives lost do not in themselves incur opportunity costs, citizens may be willing to pay to avoid them.

**Ease of identification and measurement**

Costs become increasingly more difficult to identify and measure as we progress from direct medical to direct non-medical to indirect and finally intangible. See Figure 24.

*Figure 24. Ease of identification and measurement by cost classifications*

**Health related effects**

Desired health effects, outcomes, and benefits are equivalent terms but this document use effects. Relevant health effects could include the following; mortality, morbidity, patients successfully treated, quality of life. The change in desired health effects as a result of a health care intervention could be either positive or negative. Summary of Steps in estimating costs
and effects are summarised in Table 7. The following are the suggested steps in estimating costs and effects/benefits. The first stage is investigative and qualitative and identifies the costs that are relevant and should be measured. This phase must be completed before the quantification of costs can begin.

**Table 7. Summary of the suggested steps in estimating costs and effects**

<table>
<thead>
<tr>
<th>Qualitative phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify the relevant analytic perspectives and stakeholders</td>
</tr>
<tr>
<td>Identify resources and health effects that will be changed by the intervention/policy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantitative phases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantify changes in resource utilisations and changes in health effects (both in volume units)</td>
</tr>
<tr>
<td>Estimate unit costs, and unit values</td>
</tr>
<tr>
<td>Convert resource utilisations and changes in effects into dollar values</td>
</tr>
<tr>
<td>Conduct sensitivity analysis</td>
</tr>
</tbody>
</table>

**Qualitative phase in estimating costs and effects**

Defining the perspectives and the boundaries of the study will enable the relevant resource utilisations and health effects to be isolated. The range of costs that will be altered by the event or decision and the way in which they are valued may be different for different perspectives. Costs that are applicable to a health care provider may differ from those considered relevant by a patient or by a funding agency. The opportunity cost of an option or event may differ between perspectives because different decision-makers may have different alternatives available to them, and a particular option may not be valued in the same way by all. A number of different perspectives may be relevant and frequently stakeholder and expert surveys may be necessary to investigate and delineate the relevant perspective/s. It is also important to define duration of the study because costs may occur at different times in the future and discounting may be required. All resources that will be altered or influenced by the intervention should be identified and listed.

**Quantitative phase in estimating costs and effects**

Resources consumed should first be quantified in physical units such as the number of x-rays or patients treated. Unit costs may then be estimated from statistical or accounting records. Hospital unit costs are frequently recorded and measured using diagnosis-related groups (DRGs). DRGs group cases according to the diagnosis and the hospital resources these cases consume. They are in effect hospital unit costs for different case categories. The cost of a resource is the quantity of resource units consumed multiplied by the unit cost of the resource item. Cost ($) = \sum [unit cost ($) x resource utilisation (units)]. For example, the cost of x-rays is the unit cost of an x-ray multiplied by the number of x-rays. Effects are measured in clinical or quality of life units and may be converted into dollar values of benefits by multiplying the volumes effects by unit values.
Costs and effects should always be measured incrementally. Incremental analysis considers only costs and effects that will change as a result of a decision, an intervention or event. Incremental costs of an alternative are the additional costs incurred less the costs avoided (cost offsets). Costs that will not change, and those already incurred (sunk costs), should be ignored. Similarly, a benefit that has already been incurred and will not change by the event is not relevant.

Finally, sensitivity analysis should be conducted to determine the impact of different assumptions in the values of parameters and assumptions on the results of the investigation. This is essential because most economic evaluations relate to implementation of options in the future. Univariate sensitivity analysis involves changing the value of one parameter at a time holding all other parameters constant. Multivariate sensitivity analysis, frequently employs Monte Carlo analysis modelling, and simultaneously changes a set of parameters. A good overview of sensitivity analysis (including a useful glossary) from the perspective of the National Institute for Health and Clinical Excellence (NICE) is provided by (Andronis, Barton, & Bryan, 2009).

Questions to consider

- Why is it useful to delineate health care costs into direct and indirect costs?
- Why are direct costs easier to isolate and measure compared with indirect and intangible costs?
- Why are intangible costs classified as costs when no opportunity cost or market can be defined?
- Is it possible to have a negative effect?
- Is it possible to have a negative cost?
Chapter 8: Economic Evaluation in Health

Objectives and key points

The objectives of this chapter are to integrate the topics covered earlier in this document, to elucidate the theoretical foundations of economic evaluation and summarise the process involved in the economic evaluation of changes in policies and treatment options. The spectrum of economic evaluation types their uses and objectives is also considered. At the end of this chapter you should:

- Understand the theoretical basis of economic evaluation.
- Appreciate the different ways an economic evaluation can be executed.
- Recognise that the data available and research objectives dictate the type of study.
- Understand and be able to execute the steps to be taken in an economic evaluation.
- Be aware of the role of modelling as a policy evaluation tool.
- Be able to competently plan and assess economic evaluation proposals and reports.

Theoretical foundations of economic evaluation

The primary objective of economic evaluation is to investigate information on allocative efficiency. However, in practice, selecting the “best” option for society is not necessarily the most efficient. Most choices based on efficiency criteria and maximising social welfare will result in winners and losers. Policymakers must take account of the equity consequences of different courses of action and consider the need for making tradeoffs between efficiency and equity.

Refer to chapter 5 to gain additional background on:

- Economic efficiency.
- Measurement of welfare, the Pareto concept and the Kaldor-Hicks criterion to evaluate potential welfare gains.
- Equity.
- Equity/efficiency tradeoffs.

The health economic evaluation process

Drummond defines economic evaluation “as the comparative analysis of alternative courses of action in terms of both their costs and consequences” (Michael F Drummond et al., 2005, p9). Economic evaluation applied to health care (McFarland, 2014) is a systematic approach to gathering and analysing information about changes in costs and effects of different health care programmes, policy changes or treatment options. The efficiency of use of inputs (resources) in the production of health-related outcomes is investigated and the
results provide decision makers with knowledge about potential improvements in resource usage.

Economic evaluation compares changes in the opportunity cost of resources utilised with changes in the desired health effects produced (Figure 25).

![Figure 25. Economic evaluation of a health care intervention](image)

**Types of economic evaluation in health care**

Economic evaluations in health care may be split into two main groups; partial evaluations and full economic evaluations. Partial evaluations consider costs, or effects/benefits but not both. An example of a partial evaluation is a burden of illness study.

Full economic evaluations may be classified into four types according to the way in which the differences in benefits/health effects/outcomes are quantified. Resource costs are quantified in money units in all methods. These four methods are listed in Table 8.

<table>
<thead>
<tr>
<th>Analysis method</th>
<th>Health related effects ((\Delta E))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost minimisation (CMA)</td>
<td>Zero: effects are identical</td>
</tr>
<tr>
<td>Cost effective analysis (CEA)</td>
<td>Physical or clinical units</td>
</tr>
<tr>
<td>Cost utility analysis (CUA)</td>
<td>Quality of life units</td>
</tr>
<tr>
<td>Cost benefit analysis (CBA)</td>
<td>Money values</td>
</tr>
</tbody>
</table>

CMA assumes that incremental effects are identical between alternatives thus effects do not need to be quantified. The option with the lowest cost is most efficient. Cost effectiveness analysis compares differences in cost and differences and effects between two health care interventions. CEA is the most common method used in health economics. Cost utility analysis compares differences in cost with differences in effects quantified in quality of life units. The choice criteria for CEA and CUA is the incremental cost effectiveness ratio, that is, \(\Delta C/\Delta E\). A lower cost effectiveness ratio indicates greater efficiency and is favoured above a higher cost effectiveness ratio.

CBA compares differences in costs and differences in effects measured in money values. Because a societal viewpoint is adopted CBA is capable of comparing health care interventions or a health care intervention compared with other potential societal investments. CBA could evaluate the incremental costs and incremental benefits of vaccinations for influenza with constructing a new highway. The best project is the one that generates the greatest net benefits.
Health Economics for Health Professionals an Aotearoa / New Zealand Perceptive  Page 63

(benefits less costs). Cost benefit analysis is rarely used in evaluating health care interventions but it is used extensively to evaluate government spending on other projects such as building a new highway or flood control of a river.

The types of health economic evaluations are summarised in Table 9.

Table 9. A summary of the main types of economic evaluation

<table>
<thead>
<tr>
<th>Evaluation technique</th>
<th>Perspective</th>
<th>Objective, to investigate:</th>
<th>Cost and effects (outcomes) measured</th>
<th>Comparisons between alternatives</th>
<th>Type of economic efficiency evaluated</th>
<th>Efficiency choice criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Partial evaluation (costs or effects only and no comparator)</td>
<td>Outcome analysis</td>
<td>Patient</td>
<td>Health related effects</td>
<td>Effects only</td>
<td>No</td>
<td>None</td>
</tr>
<tr>
<td>B. Full Economic evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.1 Cost effectiveness evaluation and its variants (costs measured in money units, effects quantified in other ways)</td>
<td>Cost-minimisation analysis (CMI)</td>
<td>Clinical manager, practice manager or hospital manager.</td>
<td>Allocation of resources between treatments for the same condition where effects are identical in all respects.</td>
<td>Costs only (usually direct medical).</td>
<td>Yes</td>
<td>Productive (Technical)</td>
</tr>
<tr>
<td></td>
<td>Cost-effectiveness analysis (CEA)</td>
<td>Clinical manager, practice manager or hospital manager.</td>
<td>Allocation of resources between treatments for the same condition. Effects are in the same dimension but achieved to differing degrees.</td>
<td>Both: compares direct medical, total direct, or direct plus indirect costs and effects.</td>
<td>Yes</td>
<td>Productive (Technical)</td>
</tr>
<tr>
<td></td>
<td>Cost-utility analysis (CUA)</td>
<td>Health funder/purchaser and Ministry of Health for costs. Patient for utility measure.</td>
<td>Allocation of resources across different treatments over a range of conditions within the health sector. Maximise health benefit for a defined population from a given budget.</td>
<td>Both: compares direct medical costs and effects.</td>
<td>Yes</td>
<td>Partly productive (Technical) and partly allocative efficiency.</td>
</tr>
<tr>
<td>B. 2 Cost-benefit analysis (CBA) (costs and effects measured in money units)</td>
<td>Society</td>
<td>Allocation of resources between different uses across the whole economy to maximise societal welfare.</td>
<td>Both: compares costs and benefits.</td>
<td>Yes</td>
<td>Allocative efficiency</td>
<td>Net present benefits</td>
</tr>
</tbody>
</table>

Decision-making

Economic evaluation criteria are merely one set of factors in decision-making and most decisions require a tradeoff of additional cost against additional effectiveness. The choices and interpretation of the incremental cost effectiveness ratio may be illustrated by means of the cost effectiveness plane where incremental cost of the new intervention compared with the existing intervention is plotted on the vertical axis and incremental effect
on the horizontal axis. The graph of incremental cost versus incremental effect may then be analysed with respect to the four quadrants in the graph. With any cost effectiveness analysis four outcomes are possible (Black, 1990).

Potential difficulties of interpreting the results of cost effective analysis is illustrated by the cost effectiveness plane. With any cost effectiveness analysis four outcomes are possible, two of which will involve tradeoffs (Figure 26).

![Cost Effectiveness Plane](image1)

*Figure 26. The cost effectiveness plane*

![Cost Effectiveness Plane Simulation](image2)

*Figure 27. Example of a Monte Carlo simulation of the cost effectiveness plane*
Figure 27 presents the results of a Monte Carlo simulation of the cost effectiveness plane from a study on obstructive sleep apnoea syndrome (H. M. Scott, Scott, Mihaere, & Gander, 2007).

Health economic modelling

Models replicate relevant aspects (but not all) of reality and are used to simulate different alternative choices and assumptions. They may be categorised as deterministic and stochastic, decision analytic models. Stochastic models (also called statistical models) incorporate random variables and produce different outcomes each time they are run. Monte Carlo simulation involves running a stochastic model multiple times to approximate the probability of the model’s outcomes occurring. They may be used to evaluate uncertainty in stochastic models. Deterministic models produce the same results each time the model is implemented. Economic modelling may be applied to all types of economic evaluation.

An example of a simple spreadsheet CEA model comparing two treatment options could have input cells containing resource volumes consumed, unit costs, effects produced; and output cells delivering total incremental cost and effects [(ΔC) and (ΔE)] and the incremental cost effective ratio (ΔC/ΔE). The user may conduct sensitivity analysis by altering input values and noting how the outputs change.

Check lists for assessing economic evaluations

The literature provides a number of suggested methods and the evaluation process may be described slightly differently by different authors, for example; (Michael F Drummond et al., 2005; Gold, Siegel, Russell, & Weinstein; Husereau et al.; Massetti et al., 2015; PHARMAC; Rabarison, Bish, Massoudi, & Giles, 2015; The National Institute for Health and Care Excellence (NICE), 2014).

The economic evaluation process

A common set of steps is applicable to all types of economic evaluations. These steps are summarised below. This process is set out in linear form but some steps may be combined and there will often be feedback loops.

1. Problem statement
2. Perspective/s
3. Constraints to implementation
4. Comparators
5. Identify resources and effects that would change
6. Evaluation method (CBA, CUA, CEA, CMA)
7. Quantify volume changes in resource utilisations and effects
8. Establish unit costs/values
9. Convert volume changes in resource utilisations into $ costs and if CBA convert effects into $ benefits
10. Time horizon and discounting
11. Calculate relevant choice criteria
12. Sensitivity analysis
13. Interpretation and recommendations
14. Monitor implementation
15. Make adjustments to the intervention if necessary

Some of the steps in this process will be discussed in greater detail and others were discussed in chapter 7.

Problem statement: The problem should be clearly defined. Frequently, only the tip of the problem will be evident with the bulk submerged and unseen (the iceberg phenomena) until some preliminary qualitative investigation is undertaken. Some background information will need to be gathered to clearly define the problem and set the aims and objectives. It is often useful to encapsulate the aims in the form of a one sentence research question and then list the specific objectives arising out of this.

Constraints to implementation: Before commencing an economic evaluation the relevant technical or clinical evidence and any constraints to implementation should be considered. If the proposed intervention could not be implemented because of ethical or cultural restraints or if the intervention is ineffective then the process of evaluation should stop at this point.

Comparators: All relevant alternatives should be identified to ensure that all options are considered. Usually, relevant comparator will for the intervention or treatment that will be replaced by the proposed intervention. If there are no alternative options, then the relevant comparator will be the status quo or no intervention. Doing nothing is a valid comparator.

Identify resources and effects: Identify resources and effects that would be changed by the intervention or policy. Quantify the changes in the volume of resources utilised and health effects produced. It is important to estimate changes in resource and effects in natural units, for example, hospital bed days utilised or quality adjusted life years gained to facilitate updating the study to reflect price changes for resources and effects.

Evaluation method: Selecting an evaluation method will be dictated by the study objectives, the perspective/s and the way in which the information from the investigation will be used for decision-making and data availability. Full evaluation methods could include; CBA, CUA, CEA. Partial evaluations address either health effects or costs. The most common technique is CEA, while CUA and CBA are the least frequent.
Establish $ unit costs/ unit values: Unit costs and unit values may be derived from recorded health data and from the literature. Costs and benefits: Convert changes in resource utilisations into changes in costs, and if conducting a cost benefit analysis, convert changes in effects into dollar benefits. The cost of a resource is the quantity of resource units consumed multiplied by the unit cost of the resource item. Cost ($) = \sum \text{unit cost ($) x resource utilisation (units)}. For example, the cost of x-rays is the unit cost of an x-ray multiplied by the number of x-rays. The benefits of an intervention are the numbers of units of effectiveness produced multiplied by the unit value. For example, the benefits of saving a life equals the statistical value of the human life multiplied by the numbers of lives saved. Costs and effects that cannot be quantified should be described in detail.

Time horizon and discounting: As future costs and effects are not valued as high as more immediate costs and effects. All changes in costs and effects must be discounted to a common time period. It is important to define the time horizon and the need for discounting because costs and effects may occur at different times. Often the costs of a new intervention are incurred before any benefits are derived. PHARMAC (PHARMAC, 2012) recommend discounting all base case costs and benefits in CUAs at a 3.5% and use rates of 0% and 5% in the sensitivity analyses.

Calculate relevant choice criteria: Choice criteria include cost effectiveness ratios and net benefits if conducting a CBA. Sensitivity analysis: Economic evaluation process investigates future likely events, accordingly, sensitivity analysis is required to take account of uncertainty in the estimates of the parameters. In addition to the base case, plausible ranges of parameters should also be investigated. Sensitivity analysis may be univariate where one parameter is changed holding all others constant or multi variate where all parameters are simultaneously altered.

Interpretation: Consider the potential changes in efficiency and equity, draw conclusions and select the option that best meets the study objectives. Investigate tradeoffs that may be necessary and select the best option/s: The most efficient options have the lowest incremental cost effective ratio \(\Delta C/\Delta E\) or the largest net benefit. However, as economic evaluation provides information on economic efficiency the results cannot in itself provide the best solution. Although societies aggregated welfare may be maximised by an efficient intervention there will almost always be winners and losers within society, Decision-makers will need to consider other factors in addition to the economic efficiency results of the investigation.

Monitoring: Economic evaluation focuses on the future which is uncertain, accordingly adjustments may need to be made when the intervention is implemented.
Books and Websites you may find useful

Any first-year economics text, for example, (Stewart, 2012) or (Gans, 2015), would be useful for obtaining additional information on basic economic concepts.

For the health economics component (McPake, Normand, & Smith, 2013), (M. G. Myriam Hunink, 2014) and the websites listed would be of value. https://www.pharmac.govt.nz

References


The author

Dr Guy Scott has a Masters in Agricultural Science (Specialising in Economics) from the University of Canterbury and a PhD in Public Policy from Victoria of Wellington (Thesis title: Economic Evaluation Applied to Health Policy in New Zealand).

He was a Senior Lecturer in the School of Economics and Finance at Massey University. He is Kāi Tahu, a Te Au Rangahau Research Associate, and is a Fellow of Te Mata o te Tau (Academy for Māori Research and Scholarship). Guy is a foundation member of the International Health Economics Association (iHEA). He is widely published in worldwide journals and is a frequent presenter at international conferences.
Health economics for health professionals: An Aotearoa/New Zealand perspective

Scott, Guy

2016

http://hdl.handle.net/10179/9831
11/05/2020 - Downloaded from MASSEY RESEARCH ONLINE