AFRICAN ECONOMIC RESEARCH CONSORTIUM (AERC)

COLLABORATIVE MASTERS DEGREE PROGRAMME (CMAP) IN ECONOMICS FOR SUB-SAHARAN AFRICA

JOINT FACILITY FOR ELECTIVES

Teaching Module Materials

ECON 531 - Health Economics I

(Revised: August, 2020)

Facebook   Twitter   Website   Email   Website

Table of Contents

Page

1   Introduction to Health and Healthcare

Learning Objectives

After this introductory lecture to epidemiology, you should be able to:

(a)   Distinguish between public health; community health and population health;

(b)   Define epidemiology;

(c)   Demonstrate understanding of the role of epidemiology in public health;

(d)   Summarize the historical evolution of epidemiology;

(e)   Describe the elements of a case definition and state the effect of changing the value of any of the elements;

(f)   List the key features and uses of descriptive epidemiology;

(g)   List the key features and uses of analytic epidemiology;

(h)   List the three components of the epidemiologic triad;

(i)   List and describe primary applications of epidemiology in public health practice; and

(j)   List and describe the different modes of transmission of a communicable disease in a population.

Introduction

People who practice public health come from many walks of life, for example: nurses and doctors running screening programmes; local residents campaigning for better housing; engineers drilling bore holes to provide clean water for villagers in a developing country; politicians introducing legislation to ban cigarette smoking in public places; and 'pop stars' who speak to school children on the dangers of drug misuse. These are a few examples. Many people will not identify their activities as 'public health'. What links these activities is improving the health of populations or communities. Such a broad range of activities illustrates that the factors which influence health are complex and wide-ranging. Any attempt to understand and change them

must involve many disciplines, and the study of public health draws on the expertise of people from a variety of backgrounds. Statistics, psychology, sociology, microbiology,

politics and management are some of the specialties which contribute to the study of public health. Nevertheless, epidemiology has a central role (Carr et al., 2007).[1]

Public health is the science and art of preventing disease, prolonging life as well as promoting health through organized community efforts for the sanitation of the environment, the control of community infections, the education of the individual in personal health, the organization of medical and nursing g services for the early diagnosis and preventive treatment of disease, and the development of the social machinery which will ensure that every individual in the community attains a standard of living that is adequate for the maintenance or improvement of health (Evans Jr., n.d).[2]

Further Reading
#1 Creatively listen to Introduction to Public Health available at: https://www.youtube.com/watch?v=-dmJSLNgjxo

#2 Starfield, B. (2001) Basic concepts in population health and health care, *J Epidemiol Community Health 2001;55:452–454.* Available at: https://jech.bmj.com/content/jech/55/7/452.full.pdf

Community health is generally rooted in the collective efforts of individuals and organizations who work to promote health within a geographically or culturally defined group. Community health initiatives function as "multi-sector and multi-disciplinary collaborative enterprises" that use evidence-based strategies to "engage and work with communities, in a culturally appropriate manner." The progress and success of these initiatives originate from the community members, who are collectively empowered to address self-identified vulnerabilities (e.g., education, employment, public safety). In other words, the community and its relevant characteristics are—in and of themselves—

---

[1] Carr, S., N. Unwin and T. Pless-Mulloli (2007) *An Introduction to Public Health and Epidemiology*, Second Edition, (Berkshire: Open University Press). Available at: https://alraziuni.edu.ye/uploads/pdf/book1/nursing/An%20Introduction%20to%20Public%20Health%20and%20Epidemiology%20(2007).pdf
[2] Evans Jr. (n.d) Basic concepts in public health, available at: https://samples.jbpub.com/9780763758226/58226_ch02_final.pdf

1

considered to be "an essential determinant of health" for each individual who is part of, or becomes affiliated with, a community's given membership (Marracini et al., 2017).[3]

Population health is a relatively new term and does not yet have an agreed-upon definition. Whether population health refers to a concept of health or a field of study of health determinants is not clear, and there is debate, often heated, about whether population health and public health are identical or different (Kindig, 2007).[4] In addition, discussions of population health use many terms, such as outcomes, disparities, determinants, and risk factors. These terms are often used imprecisely, particularly when different disciplines, such as medicine, epidemiology, economics, and sociology, are involved. Although these are sometimes merely minor semantic differences, the meanings are often unclear and cause significant miscommunication.

Maraccini et al. (2017) argue that population health uses an outcome-driven approach to "manage" health for a specific group of individuals, typically defined by attribution. These interventions involve the tracking and measurement of "health status indicators" (e.g., high blood pressure, cholesterol) within these groups to provide insight and direction on how to best prevent the onset or future development of certain health conditions (e.g., ischemic cardiac disease). Health determinants, such as healthcare access, genetics, and individual behavior, also tend to be included in this description, as they play an influential role in an individual's history and current health status. Given their nature, the majority of population health interventions tend to be led by healthcare organizations, including Accountable Care Organizations, who have a responsibility—financial or otherwise—to report outcomes involving the patients under their management or care.

---

[3] Maraccini, A. M., P. Galiatsatos, M. Harper et al. (2017) Creating clarity: Distinguishing between community and population health, *The American Journal of Accountable Care, Vol.6, No. 17, pp.32-37*.
[4] Kindig, D. (2007) Understanding Population Health Terminology, *The Milbank Quarterly, Vol. 85, No. 1, 2007, pp. 139–161*.

2

Table 1.1 provides a list of dimensions (i.e., key descriptors, populations targeted, interventions utilized, and measurement) to further highlight areas of distinction and overlap between community health and population health.

Table 1.1 Areas of Distinction and Overlap in Community and Population Health

Dimensions	Community Health	Community and Population Health	Population Health
Key Descriptors	Collaborative, initiatives, empowerment	Disease prevention, health promotion	Outcomes, management, accountability
Populations Targeted	Geographically and/or culturally defined	Patient population within a community-defined area	Typically defined by attribution
Interventions Utilized	Broad scale (social determinants)	Disease prevention, health promotion	Population-specific (access; prevention)
Measurement	Overall health; community efforts	Community health needs assessments	Health status indicators; healthcare outcomes

Source: Maraccini et al. (2017)

Further Reading
Kindig, D. and G. Stoddart (2003) What Is Population Health? *American Journal of Public Health | March 2003, Vol 93, No. 3, pp. 380-383*.

Kindig, D. (2007) Understanding Population Health Terminology, *The Milbank Quarterly, Vol. 85, No. 1, 2007, pp. 139–161*. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2690307/pdf/milq0085-0139.pdf

Maraccini, A. M., P. Galiatsatos, M. Harper et al. (2017) Creating clarity: Distinguishing between community and population health, *The American Journal of Accountable Care, Vol.6, No. 17, pp.32-37*.

The term '**epidemiology'** comes from the following three Greek words:- (i) **epi**, meaning "on or upon," (ii) **demos**, meaning "people," and (iii) **logos**, meaning "the study of." Myriad definitions have been proposed. However, the following definition captures the underlying principles and the public health spirit of epidemiology. Dicker (n.d)[5] defines epidemiology as the **study** of the **distribution** and **determinants of health-related**

---

[5] Dicker (n.d) *Principles of Epidemiology*, Second Edition, available at:
http://www.ciphi.ca/hamilton/Content/documents/principles.pdf

states or events in **specified populations**, and the **application** of this study to the control of health problems.

This definition of epidemiology includes several terms which reflect some of the important principles of the discipline. As you study this definition, refer to the description of these terms below.

Study. Epidemiology is a scientific discipline, sometimes called "the basic science of public health." It has, at its foundation, sound methods of scientific inquiry.

Distribution. Epidemiology is concerned with the **frequency** and **pattern** of health events in a population. Frequency includes not only the number of such events in a population, but also the rate or risk of disease in the population. The **rate** (number of events divided by size of the population) is critical to epidemiologists because it allows valid comparisons across different populations.

Activity

Brainstorm with your virtual neighbor about the meaning and differences between **incidence** and **prevalence** of a health-related event.

Prevalence

The prevalence represents existing cases of a disease and can be seen as a measure of disease status (Noorzdij et al., 2010). It is the proportion of people in a population having a disease given by the following formula:

$$\Pr{evalence} = \frac{Number\ of\ subjects\ having\ the\ disease\ at\ a\ time\ point}{Total\ number\ of\ subjects\ in\ the\ population}$$

The prevalence is often useful as it reflects the burden of a disease in a specific population. This is not limited to the burden of disease in terms of monetary costs. It

also reflects burden in terms of life expectancy, morbidity, quality of life, or other indicators. Knowledge of the burden of disease can help decision makers to determine where investments in health care should be targeted.

Incidence

While the prevalence represents the existing cases of a disease, the incidence reflects the number of new cases of disease within a certain period and can be expressed as a risk or an incidence rate.

Risk

The risk is the probability that a subject within a population will develop a given disease, or other health outcome, over a specified follow-up period. It can be calculated by dividing the number of subjects developing the disease over a certain period by the total number of subjects followed over that period:

$$Risk = \frac{Number\ of\ subjects\ developing\ the\ disease\ over\ a\ time\ period}{Total\ number\ of\ subjects\ followed\ over\ that\ time\ period}$$

This risk can be interpreted as an estimation of the risk of disease in an individual subject. However, to interpret a risk appropriately, it is necessary to know the time period to which it applies. Without the definition of a time period, the computed measure of risk is a meaningless value.

For the calculation of a risk, a few assumptions need to be made. First, because the risk reflects new cases of disease, all subjects should be free of this disease at the start of the follow-up period. Second, all subjects should be followed over the total period of time during which the risk is measured. This second requirement can lead to a few problems, especially in studies with a relatively long follow-up period. The longer the follow-up period is, the higher is the chance subjects will become lost to follow-up. In addition, subjects can drop out of the study because of causes that 'compete' with the

outcome of interest. For example, if one aims to study death of cardiovascular causes, death of any other cause, e.g. a car accident, can be considered as a competing risk. Both loss to follow-up and competing risks can lead to an underestimation of the risk, because the subjects leaving the study are not any longer able to experience the event of interest and will therefore not be able to add to the numerator in the formula.

Incidence Rate

The second measure of disease frequency that expresses incidence is the incidence rate. It can be calculated by dividing the number of subjects developing a disease by the total time at risk for all people to get the disease. The denominator of this formula includes a measure of time instead of just a number of subjects. The incidence rate should therefore be interpreted as an instantaneous concept, like speed:

$$Incidence\ rate\ = \frac{Number\ of\ subjects\ developing\ the\ disease}{Total\ time\ at\ risk\ for\ the\ disease\ for\ all\ followed\ subjects}$$

Like for the measure of risk, one assumes for the calculation of the incidence rate that all subjects are free of the disease of interest at the start of the study. However, an important advantage of the incidence rate over the risk is that it is not required for subjects to complete the total follow-up time and only the actual time at risk is taken into account.

Under conditions in which rates do not change with time (a steady state), the incidence rate can be interpreted as the reciprocal of the average time until an event occurs, also called the waiting time. When calculated over a short period of time, the risk and the incidence rate will be rather similar, because the influence of loss to follow-up and competing risks which may flaw risk will only be small.

Incidence versus Prevalence

Factors that influence the prevalence are the number of incident cases, the deaths, and the recoveries, as is depicted in Figure 1.1. Given a steady state, the prevalence approximately equals the product of the incidence rate and the mean duration of disease.
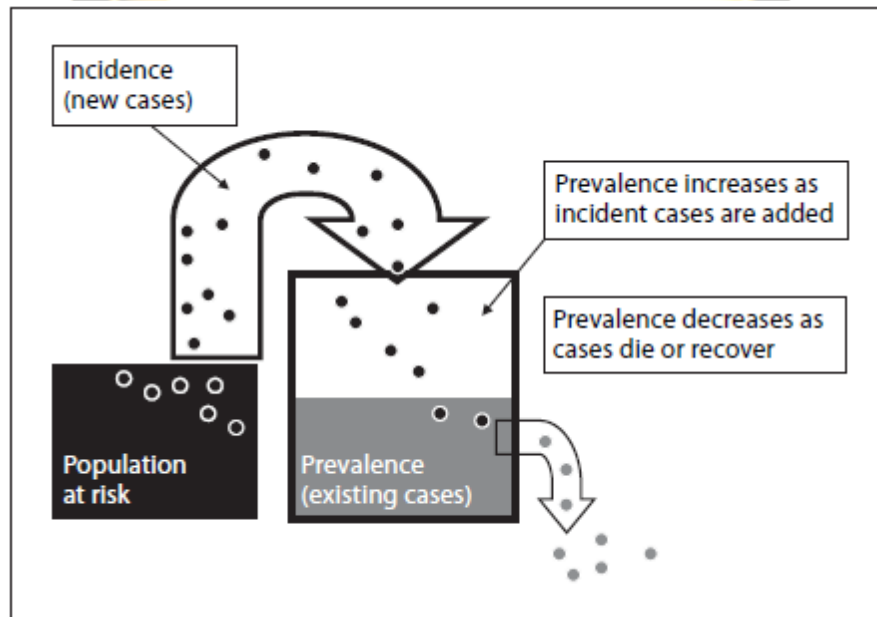
Figure 1.1 Relationship between Incidence and Prevalence

Table 1.2 presents the properties of all described measures of disease frequency are summarized as well as some of their uses.

Table 1.2 Summary of Characteristics of Measures of Disease Frequency

	Prevalence	Incidence	
Represents	Existing cases at a time point	New cases over a period	
Use	– Reflects disease burden	– Assessment of disease aetiology	
	– Can be used for planning of health care facilities	– Identification of risk factors	
		Risk	Incidence rate
Synonyms	Prevalence proportion	Cumulative incidence	Incidence density
		Incidence proportion	Hazard
Range	0–1 (0–100%)	0–1 (0–100%)	0 to infinity
Interpretation	Proportion	Probability	Reciprocal of waiting time

Source: Noorzdij et al. (2010)

Further Reading

Principles of infectious disease and epidemiology, available at: https://www.youtube.com/watch?v=xlfXRAkhvhw

Deadliest Plague of the 20th Century: Flue of 1918, available at: https://www.youtube.com/watch?v=UDY5COg2P2c

Journalist goes undercover at "wet markets", where the coronavirus started, available at: https://www.youtube.com/watch?v=Y7nZ4mw4mXw

Noorzdij, M., F. W. Dekker, C. Zoccali and K. J. Jager (2010) Measures of Disease Frequency: Prevalence and Incidence, *Nephron Clinical Practice, Vol. 115, pp. c17-c20*, available at: https://www.karger.com/Article/Pdf/286345

Pattern refers to the occurrence of health-related events by **time**, **place**, and **personal characteristics**. *Time characteristics* include annual occurrence, seasonal occurrence, and daily or even hourly occurrence during an epidemic. *Place characteristics* include geographic variation, urban-rural differences, and location of worksites or schools. *Personal characteristics* include demographic factors such as age, race, sex, marital status, and socioeconomic status, as well as behaviors and environmental exposures.

This characterization of the distribution of health-related states or events is one broad aspect of epidemiology called **descriptive epidemiology**. Descriptive epidemiology provides the *What*, *Who*, *When*, and *Where* of health-related events.

Determinants. Epidemiology is also used to search for causes and other factors that influence the occurrence of health-related events. **Analytic epidemiology** attempts to provide the *Why* and *How* of such events by comparing groups with different rates of disease occurrence and with differences in demographic characteristics, genetic or immunologic make-up, behaviors, environmental exposures, and other so-called potential risk factors. Under ideal circumstances, epidemiologic findings provide sufficient evidence to direct swift and effective public health control and prevention measures.

Health-related states or events. Originally, epidemiology was concerned with epidemics of communicable diseases. Note that epidemiology has been extended to endemic communicable[6] diseases and non-communicable[7] diseases. More recently, epidemiologic methods have been applied to chronic diseases, injuries, birth defects, maternal-child health, occupational health, and environmental health. Now, even behaviors related to health and well-being (amount of exercise, seat-belt use, etc.) are recognized as valid subjects for applying epidemiologic methods. In this lecture we use the term "disease" to refer to the range of health-related states or events.

Historical Evolution of Epidemiology

The work of Thomas Sydenham (1624-89), in identifying disease entities, and of John Graunt (1620-74) in using quantitative methods for the study of plague is generally recognized as initiating a population approach to problems of disease. After that, epidemiology struggled to achieve recognition as a scientific discipline. It was only at the turn of the nineteenth century, with chronic diseases such as cancer and cardiovascular diseases emerging as major killers, and with the development of Pearsonian and Fisherian statistics, that new epidemiological methods began to be developed and applied in a systematic way[8].

Further Reading
For a historical evolution of epidemiology, read: **Chapter 2: Historic Developments in Epidemiology**, available at: https://samples.jbpub.com/9780763766221/66221_CH02_5398.pdf

The Elements of a Case Definition

Before counting cases, the epidemiologist must decide what to count, that is, what to call a case. For that, the epidemiologist uses a case definition. A case definition is a set of standard criteria for classifying whether a person has a particular disease, syndrome,

---

[6] Or infectious or contagious or transmissible

[7] Or noninfectious or non-contagious or non-transmissible

[8] https://grantome.com/grant/NIH/G13-LM010884-02

or other health condition. Some case definitions, particularly those used for national surveillance, have been developed and adopted as national standards that ensure comparability. Use of an agreed-upon standard case definition ensures that every case is equivalent, regardless of when or where it occurred, or who identified it. Furthermore, the number of cases or rate of disease identified in one time or place can be compared with the number or rate from another time or place[9].

Further Reading

Timen, A., G. J. J. van Doornum, M. Schutten et al. (2006) Public health implications of using various case definitions in The Netherlands during the worldwide SARS outbreak, *Clinical Microbiology and Infection, Vol. 12, No. 12, pp. 1214–1220*.

Descriptive Epidemiology

Descriptive epidemiology refers to epidemiological studies and activities with descriptive components that are much stronger than their analytic components or that fall within the descriptive area of the descriptive-analytic spectrum. Descriptive epidemiology deals with the occurrence of disease, in terms of both geographical comparisons and descriptions of temporal trends. Descriptive studies are positioned at the base of the hierarchy of scientific evidence. Nonetheless their importance as the basic roots of the epidemiologic approach has not changed. In particular, disease prevalence and incidence data perform an essential role in both research and clinical settings (Naito, 2014).

Further Reading

Naito, M. (2014) Utilization and Application of Public Health Data in Descriptive Epidemiology, *Journal of Epidemiology, Vol. 24, No.6, pp.435-436*.

Chapter 4 Descriptive Epidemiology: Pattern of Disease- Person, Place and Time, available at: https://samples.jbpub.com/9780763754433/Chapter4.pdf

---

9

https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section5.html#:~:text=A%20case%20definition%20consists%20of,findings)%2C%20and%20other%20findings.

Analytic Epidemiology

As noted earlier, descriptive epidemiology can identify patterns among cases and in populations by time, place and person. From these observations, epidemiologists develop hypotheses about the causes of these patterns and about the factors that increase risk of disease. In other words, epidemiologists can use descriptive epidemiology to generate hypotheses, but only rarely to test those hypotheses. For that, epidemiologists must turn to analytic epidemiology.[10]

Causation of ill health is difficult to determine. Apart from infectious diseases, most evidence relates to the risk associated with particular factors, rather than the direct causes of ill health. Even in the case of infectious diseases, it is not known why certain people succumb to them when other seemingly similar people do not. Only certain types of causative mechanism are amenable to investigation by epidemiological methods and few diseases can be said to have a single cause. Epidemiological evidence regarding disease causation is mainly circumstantial. So how can the information on persons, place and time help in the understanding of the causes of disease? An important model, originally developed for use with infectious diseases, but which has been adapted for much broader use in recent years, is the *epidemiological triad*.[11]

Further Reading

Heaton, B. and T. Dietrich (2000) Analytic epidemiology and periodontal diseases, *Periodontology 2000, Vol. 58, 2012, pp.112–120*, available at: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1600-0757.2011.00419.x?casa_token=1nGTjCDVtOkAAAAA:xPR4s7ZOIDKEJuoL7Aat1InhCwfgDv1SU0LLHGcMJOTtNdnETXXwexfZBGuXPOqhLlPXlsWyhcrPi2OL

Chapter 2 Epidemiology, available at: https://samples.jblearning.com/9781284166606/9781284188479_CH02_Secured.pdf

Epidemiologic Triad or Epidemiologic Triangle

---

[10] https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section7.html
[11] https://www.open.edu/openlearn/health-sports-psychology/health/epidemiology-introduction/content-section-3

The epidemiological triad or the epidemiologic triangle represents the interaction between an agent, host or persons and environment or place within a specific time dimension (Figure 1.1). The epidemiological triad can also be applied to non-infectious diseases where the agent could be 'unhealthy behaviours, unsafe practices, or unintended exposures to hazardous substances'.[12]
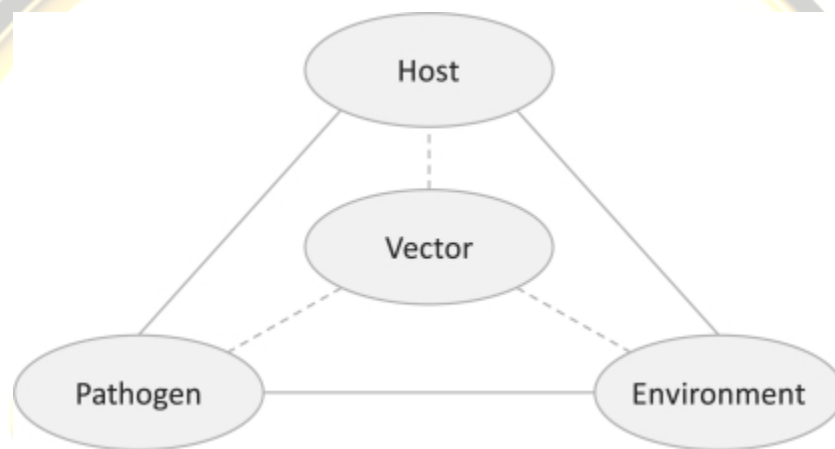
Figure 1.1 The Epidemiologic Triad

Source: https://openi.nlm.nih.gov/detailedresult?img=PMC3117366_btr226f1&req=4

Within the epidemiological triad, the main infection components are host, agent and environment. The vector is frequently related to all components making it a hub node in the transmission network, and hence a good target for infection control approaches. The agent is known as a 'necessary' factor. It has to be present for morbidity, although it may not inevitably lead to disease.[13] For the disease to occur it needs the combination of what have been called 'sufficient' factors. These would include a host, which might be an individual or group of individuals who are susceptible to the agent. Susceptibility might be on the basis of age, sex, ethnic group or occupation. Environmental factors can also be sufficient factors that combine with the agent.

---

[12] https://www.open.edu/openlearn/health-sports-psychology/health/epidemiology-introduction/content-section-3.1
[13] https://www.open.edu/openlearn/health-sports-psychology/health/epidemiology-introduction/content-section-3.1

Further Reading

Epidemiologic Triangle, available at: https://www.youtube.com/watch?v=1r_u5_p-WQA
Chapter 12 Epidemiologic Triad, available at: https://courses.lumenlearning.com/suny-buffalo-environmentalhealth/chapter/epidemiological-triad-overview/
Chapter 1 Foundations of Epidemiology, available at:
https://samples.jbpub.com/9780763766221/66221_CH01_5398.pdf

Poundstone, K. E., S. A. Strathdee and D. D. Celentano (2004) The Social Epidemiology of Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome, *Epidemiologic Reviews, Vol. 26, pp. 22-35.* (Available at: https://academic.oup.com/epirev/article/26/1/22/384205).

Principles of Epidemiology, available at:
http://www.ciphi.ca/hamilton/Content/documents/principles.pdf

Primary Applications of Epidemiology in Public Health Practice

Epidemiology is the study of the health of human populations. Its functions include but are not limited to: (a) discovering the agent, host, and environmental factors that affect health to provide the scientific basis for the prevention of disease and injury and the promotion of health; (b) determining the relative importance of causes of illness, disability, and death to establish priorities for research and action; (c) identifying those sections of the population that have the greatest risk from specific causes of ill health so that the indicated action may be directed appropriately; and (d) evaluating the effectiveness of health programs and services in improving the health of the population (Haveman-Nies et al., 2010).[14]

Further Reading

Chapter 1 Foundations of Epidemiology, available at:
https://samples.jbpub.com/9780763766221/66221_CH01_5398.pdf

Epidemiology in Public Health Practice, available at:
https://academic.oup.com/aje/article/174/7/871/116502

Díaz-Vélez et al. (2013) Clinical Epidemiology and its Relevance for Public Health in Developing Countries, available at: https://www.intechopen.com/books/current-topics-in-

---

[14] Epidemiology in Public Health Practice, available at: https://academic.oup.com/aje/article/174/7/871/116502

13

Different Modes of Transmission of a Communicable Disease in a Population

Communicable diseases can be transmitted by both direct and indirect modes. Direct transmission refers to the transfer of an infectious agent from an infected host to a new host, without the need for intermediates such as air, food, water or other animals. Direct modes of transmission can occur in two main ways: (a) Person to person: The infectious agent is spread by direct contact between people through touching, biting, kissing, sexual intercourse or direct projection of respiratory droplets into another person's nose or mouth during coughing, sneezing or talking. A familiar example is the transmission of HIV from an infected person to others through sexual intercourse. (b) Transplacental transmission: This refers to the transmission of an infectious agent from a pregnant woman to her fetus through the placenta. An example is mother-to-child transmission (MTCT) of HIV.[15]

Conversely, indirect transmission is when infectious agents are transmitted to new hosts through intermediates such as air, food, water, objects or substances in the environment, or other animals. Indirect transmission has three subtypes: (a) Airborne transmission: The infectious agent may be transmitted in dried secretions from the respiratory tract, which can remain suspended in the air for some time. For example, the infectious agent causing tuberculosis can enter a new host through airborne transmission. (b) Vehicle-borne transmission: A vehicle is any non-living substance or object that can be contaminated by an infectious agent, which then transmits it to a new host. Contamination refers to the presence of an infectious agent in or on the vehicle. (c) Vector-borne transmission: A vector is an organism, usually an *arthropod*, which transmits an infectious agent to a new host. Arthropods which act as vectors include houseflies, mosquitoes, lice and ticks. Arthropods are invertebrates (animals without

---

15

https://www.open.edu/openlearncreate/mod/oucontent/view.php?id=84&printable=1#:~:text=Direct%20m odes%20of%20transmission%20can,during%20coughing%2C%20sneezing%20or%20talking.

backbones), such as insects, which have segmented bodies and three pairs of jointed legs.[16]

Further Reading

Communicable Diseases Module: 1. Basic Concepts in the Transmission of Communicable Diseases, available at:
https://www.open.edu/openlearncreate/mod/oucontent/view.php?id=84&printable=1#:~:text=Direct%20modes%20of%20transmission%20can,during%20coughing%2C%20sneezing%20or%20talking.

Centre for Health Protection-Department of Health (2019) Guidelines on Prevention of Communicable Diseases in Residential Care Homes for Persons with Disabilities, available   at: https://www.chp.gov.hk/files/pdf/guideline_prevention_of_communicable_diseases_rchd.pdf

---

16

https://www.open.edu/openlearncreate/mod/oucontent/view.php?id=84&printable=1#:~:text=Direct%20modes%20of%20transmission%20can,during%20coughing%2C%20sneezing%20or%20talking.

2　Introduction to Health Economics

Learning Objectives

At the end of this topic/unit, you should be able to:-

(a) Define and distinguish between disease, illness, sickness, and health;

(b) Demonstrate understanding of the various models of health and disease (Biomedical model; Holistic health model; Biopsychosocial Model; Ethnomedical Cultural Model; and Ecological-Transactional Model);

(c) Demonstrate understanding of the various issues considered by health economics (determinants of health (other than health care); measurement of health; the demand for health care; the supply of health care; microeconomic evaluation at the treatment level; the results of economic evaluation inform the choices of suppliers and purchasers and the creation of market equilibrium; the evaluation of the entire health system; and the evaluation of mechanisms to change the health system).

It should be noted that health economics is constantly evolving sub-discipline of economics. This is because of the intertwined nature of the global economy characterized by globalization as well as the increased use of Information Communication Technology (ICT). Climate change as well as its implications for global public health is a topical issue that increasingly finds its way into contemporary discussions of health economics. An interdisciplinary approach to health economics research is instructive given the failure of policy prescriptions from specific disciplines.

Introduction

Before this unit delves into Health Economics, it defines and distinguished between health; wellness and illness. Thereafter, it considers various models of health and disease. Kernick (2003)[17] notes that against a background of increasing demands on

---

[17] Kernick, D. P. (2003) Introduction to health economics for the medical practitioner, *Postgraduate Medical Journal, Vol. 79, pp.147–150.*

limited resources, health economics is exerting an influence on decision making at all levels of health care. Health economics seeks to facilitate decision making by offering an explicit decision making framework based on the principle of efficiency. It is not the only consideration but it is an important one and practitioners will need to have an understanding of its basic principles and how it can impact on clinical decision making.

Disease, Illness, Sickness, and Health

Boyd (2000)[18] presents various definitions of disease, illness, sickness, health, healing and wholeness. He notes that disease is a pathological process, most often physical as in throat infection, or cancer of the bronchus, sometimes undetermined in origin, as in schizophrenia. The quality which identifies disease is some deviation from a biological norm. There is objectivity about disease whereby doctors are able to see, touch, measure, as well as smell. Diseases are valued as the central facts in the medical view.

Illness is a feeling, an experience of being unhealthy which is entirely personal, interior to the person of the patient. Oftentimes it accompanies disease, but the disease may be undeclared, for example in the early stages of cancer or tuberculosis or diabetes. Sometimes illness exists where no disease can be found. Traditional medical education has made the deafening silence of illness-in-the-absence-of-disease unbearable to the clinician. The patient can offer the doctor nothing to satisfy his senses (Boyd, 2000).

Sickness is the external and public mode of unhealth. Sickness is a social role, a status, a negotiated position in the world, a bargain struck between the person henceforward called 'sick', and a society which is prepared to recognize and sustain him. The security of this role depends on a number of factors, not least the possession of that much treasured gift, the disease. Sickness based on illness alone is a most uncertain status. But even the possession of disease does not guarantee equity in sickness. Those with a chronic disease are much less secure than those with an acute one; those with a psychiatric disease than those with a surgical one. Best is an acute physical disease in

---

[18] Boyd, K. M. (2000) Disease, illness, sickness, health, healing and wholeness: exploring some elusive concepts, *Journal of Medical Ethics: Medical Humanities, Vol. 26, pp.9–17*.

a young man quickly determined by recovery or death—either will do, both are equally regarded (Boyd, 2000).

The modern understanding of health became official when the World Health Organization (WHO), at the time of its establishment in 1948, included the definition of health in its Constitution. This generally accepted definition states that "**health** is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity". This definition promoted for the first time that, in addition to physical and mental health, social welfare is an integral component of the overall health, because health is closely linked to the social environment and living and working conditions (Svalastog et al., 2017).[19]

Further Reading

Boyd, K. M. (2000) Disease, illness, sickness, health, healing and wholeness: exploring some elusive concepts, *Journal of Medical Ethics: Medical Humanities, Vol. 26, pp.9–17.*

Sartorius, N. (2006) The Meanings of Health and its Promotion, available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2080455/pdf/CroatMedJ_47_0662.pdf

Svalastog et al. (2017) Concepts and definitions of health and health-related values in the knowledge landscapes of the digital society, available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5778676/pdf/CroatMedJ_58_0431.pdf

Health and Illness, https://www.aihw.gov.au/getmedia/fc2c053f-7707-437e-90b3-227818bba0a5/1_1-health-illness.pdf.aspx

Boruchovitch and Mednick (2002) The meaning of health and illness: some considerations for health psychology, http://www.scielo.br/pdf/pusf/v7n2/v7n2a06.pdf

American Economic Review, Vol. 53, No. 5, pages 941-973. Available at: https://web.stanford.edu/~jay/health_class/Readings/Lecture01/arrow.pdf

---

[19] Svalastog, A. L., D. Donev, N. J. Kristoffersen et al. (2017) Concepts and definitions of health and health-related values in the knowledge landscapes of the digital society, available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5778676/pdf/CroatMedJ_58_0431.pdf

Models of Health and Illness

Myriad models of health and illness are available for use by family physicians as guides in delivering health care to their patients. A model is a belief system or plan for making sense of the puzzling phenomena confronting patients and physicians. The following models of health and illness are considered in what follows. Biomedical model; Holistic health model; Biopsychosocial Model; Ethnomedical Cultural Model; and Ecological-Transactional Model.

Biomedical Model

Traditional biomedical models, bolstered by the famous Flexner report on medical education in 1910, consolidated the professional position of physicians and produced phenomenal technological, molecular, and pharmacological advances. The biomedical model is rooted in the assumptions of mind-body dualism, biological reductionism, and linear causality. This model assumes that there exists a primary biophysical cause for each disease that is objectively identifiable by an observer who is independent of the disease process (Miller, 1998)[20]. From a simplistic viewpoint, the principal characteristics of a biomedical model derived from these assumptions are (i) physician-centered, (ii) specialist-dominated, (iii) credential-oriented, (iv) memory-based, (v) disease-oriented, (vi) technology-dominated, (vii) biological process oriented, and (viii) single case-centered. As a result, a physician using the biomedical approach explores relationships between symptoms and somatic disorders by pursuing those data that reveal anatomical or psychophysiological lesions. Treatment then consists of interventions designed to counteract the somatic disease process. In the biomedical model, health is viewed as an absence of disease. Biomedical physicians act as scientists, fighters, and technicians, and their styles tend toward paternalism.

Holistic Health Model

Over time, three new models have emerged as efforts to supplement the biomedical model. The oldest of these new models is the holistic health model. The general public,

---

[20] Miller, W.L. (1988) Models of Health, Illness, and Health Care. In: Taylor R.B. (eds) Family Medicine. Springer, New York, NY.

paraprofessionals, and allied health care providers, such as chiropractors, nutritionists, physical therapists, nurses, psychotherapists, and pastoral counselors, also have recognized problems with the biomedical model and have rallied behind a holistic health model approach (Miller, 1988). Seizing on the public perception of patient powerlessness and alienation from a high-tech, authoritarian health care system, the holistic health model emphasizes individual responsibility and personal growth. A patient (referred to as client) is an active partner in the health care team-client relationship. An anticipated goal in this model is high level wellness, a dynamic state of perceived being where body, mind, and spirit are in harmony. The holistic health model, also called "integral medicine," is based on the following four assumptions: (i) all states of health and disorder are psychosomatic, (ii) each person is a circular interaction of body, mind, and spirit, (iii) illness and disease result from adaptational failure, producing a disturbance in the self-regulation of one's social, behavioral, psychological, and physical dimensions, and (iv) illness is a creative opportunity. Based on these assumptions, health promotion, self-care, and education are key activities.

Biopsychosocial Model

The biopsychosocial model was proposed to respond to the problems associated with the biomedical model. This model has been well received in academic family medicine and other primary care specialties and has already demonstrated clinical and research utility. The biopsychosocial model expands the biomedical model by incorporating psychosocial factors into the scientific equation. Reductionism and linear causality are replaced by a systems perspective. Although this model is still disease-oriented, disease is now seen as belonging to a person who is part of a system of hierarchically arranged units from atom to cell to organ to body to family to culture to universe, linked by circular feedback (Miller, 1988). The context in which symptoms occur becomes as important as the symptoms themselves. The biopsychosocial physician's style is one of mutual participation, with prevention and lifestyle issues given very prominent consideration. Following a biopsychosocial evaluation, a disease and illness problem list

is drafted, and a comprehensive management plan is designed acknowledging the problem(s), stressors, and social support systems.

Ethnomedical Cultural Model

The biopsychosocial model has been expanded even farther to include cultural factors. The ethnomedical cultural model pays close attention to differences between disease and illness and emphasizes a patient's understanding of his illness. Ethnomedical cultural physicians uncover a patient's ethnomedical beliefs, assumptions, and expectations about body structure and function, such as body as machine vs body as plumbing vs body as balance of hot and cold (Miller, 1988). Diet and nutrition, pain, stress, family, social network, help seeking, illness classification, and lay and folk-expert healers, including root doctors and voodoo doctors among southern blacks, acupuncturists and herbalists among Chinese and Southeast Asians, espiritistas among Puerto Ricans, and curanderas among Mexican-Americans, also are defined in an ethnomedical approach.

Ecological-Transactional Model

The ecological-transactional model diagrammed in Figure 2.1 defines the context (ecology) and the content (transaction) of the clinical encounter and provides a map for guiding a family physician. The patient in this ecological perspective is viewed holistically as a union of biological and psychocultural factors that determine an individual's adaptive capacity or ability to adjust to environmental influences and changes (Miller, 1988). The microsystem refers to immediate environment and includes a patient's family health culture, significant others, work and religious life, and physical home. The ecosystem encompasses neighborhood and includes the remainder of a lay referral network and local community institutions, including the health care system. The macrosystem is concerned with a more global context, such as socioeconomic and political forces impinging on health care. At the individual level, the cultural factors of most significance include patient requests, individual placement and support as well as emergency medical services.

21
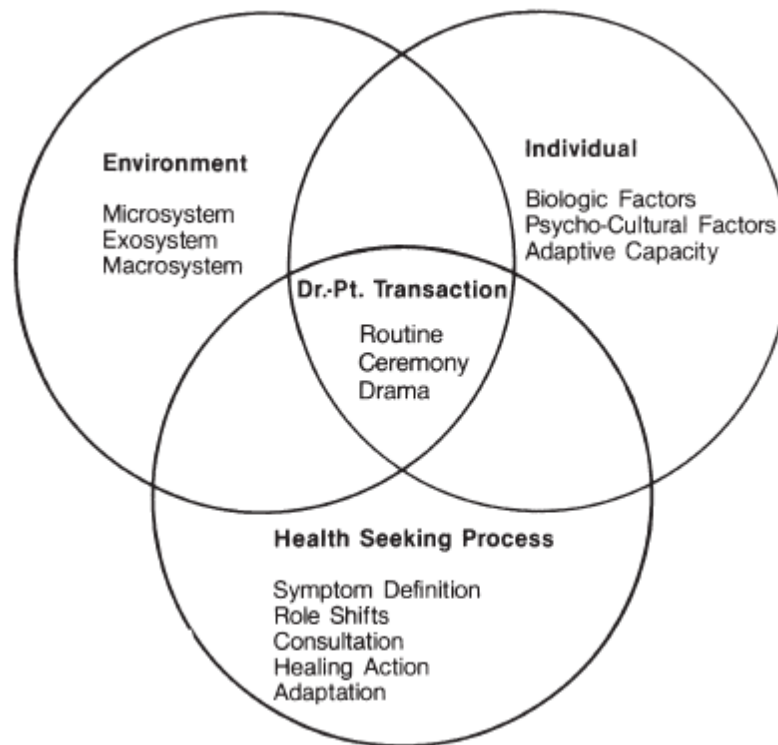
Figure 2.1 Ecological-transactional model of health, sickness, and family physician care. Source: Miller (1988)

Further Reading

Miller, W.L. (1988) Models of Health, Illness, and Health Care. In: Taylor R.B. (eds) *Family Medicine.* Springer, New York, NY.

Overview, Meaning and Scope of Health Economics

In 1963, Kenneth Arrow published "Uncertainty and the welfare economics of medical care" in The American Economic Review.[21] Savedoff (2004)[22] notes that the paper became not only one of the most widely cited articles in the field of health economics — indeed, it marked the creation of the discipline — but also a source of reference in other fields. Arrow (1963) begins with reference to the desirable properties of perfectly competitive markets, using concepts from Arrow's general welfare theorems. The article

---

[21] Arrow, K. (1963) Uncertainty and the welfare economics of medical care, *The American Economic Review, Vol. 53, pp. 941-73.*
[22] Savedoff, W.D. (2004) Kenneth Arrow and the birth of health economics, *Bulletin of the World Health Organization, Vol. 82, No. 2, pp. 139-140.*

22

then explores how the existence of "uncertainty in the incidence of disease and in the efficacy of treatment" leads competitive markets to generate an inefficient allocation of resources and contributes to the emergence of nonmarket institutions (such as trust and norms) that compensate for these market failures (Savedoff, 2004). A full discussion of the historical roots of health economics is presented in Jakovljevic and Ogura (2016).[23]

Kernick (2003)[24] has defined health economics as the discipline of economics applied to the topic of health care. Broadly defined, economics concerns how society allocates its resources among alternative uses. Scarcity of these resources provides the foundation of economic theory and from this starting point, three basic questions arise: (i) What goods and services shall we produce? (ii) How shall we produce them? And (iii) Who shall receive them? Health economics addresses these questions primarily from the perspective of efficiency— maximizing the benefits from available resources (or ensuring benefits gained exceed benefits forgone). Equity concerns are also recognized— what is a fair distribution of resources. Considerations of equity often conflict with efficiency directives. However, due to the contested nature of this area and the difficulties in quantifying equity dimensions, this element has not been a major focus of health economist's work.

Scope of Health Economics

Maynard and Kanavos (2000)[25] use Alan Williams' 'plumbing' diagram (Figure 2.2) to describe the sub-discipline of health economics.

---

[23] Jakovljevic, M. and S. Ogura (2016) Health economics at the crossroads of centuries – From the past to the future, *Frontiers in Public Health, Vol. 4, Article 115, pp.1-5*, available at: https://www.frontiersin.org/articles/10.3389/fpubh.2016.00115/full

[24] Kernick, D. P. (2003) Introduction to health economics for the medical practitioner, *Postgraduate Medical Journal, Vol. 79, pp.147–150*.

[25] Maynard, A. and P. Kanavos (2000) Health economics: an evolving paradigm, *Health Economics, Vol. 9, pp. 183–190*.
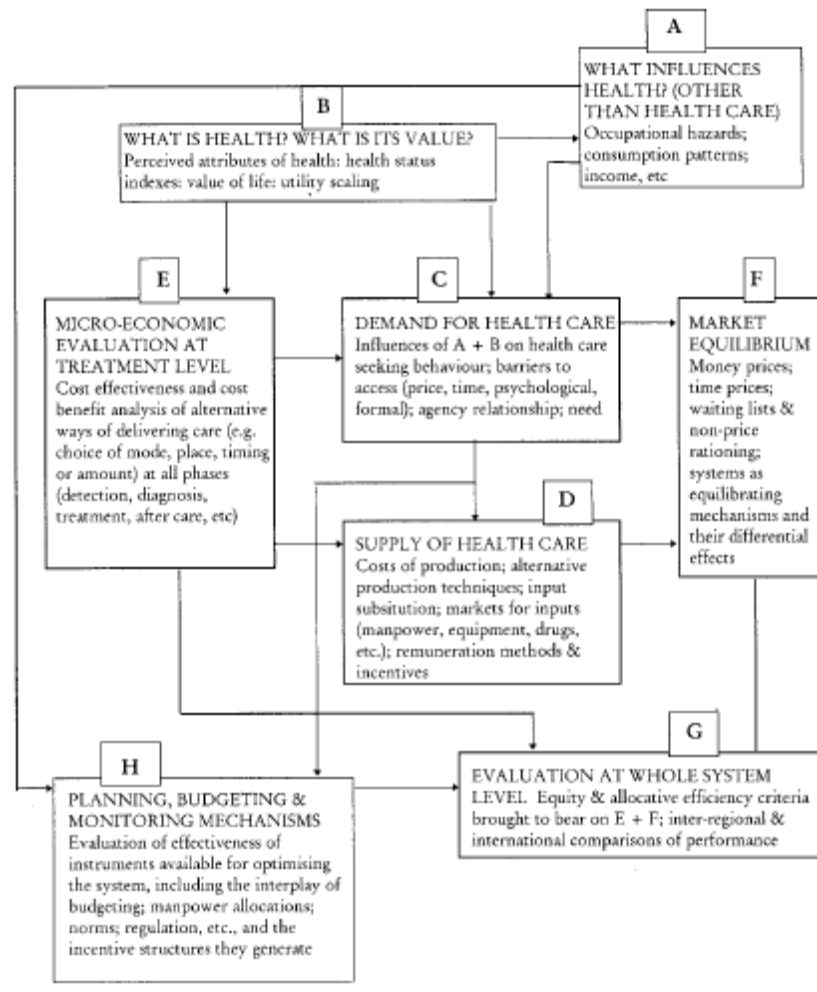
Figure 2.2 Structure of the Discipline of Health Economics. Source: Maynard and Kanavos (2000)

The complex linkages as well as the scope of health economics is described if one follows this figure from boxes A to H. Box A focuses on the determinants of health (other than health care). Whilst many have argued that the influence of the level and distribution of income and wealth, as well as education levels, housing conditions, leisure behaviours and nutritional habits, have significant effects on health, economists have carried out relatively little work on modelling and estimating such relationships.

Box B in Figure 2.2 is the measurement of health. What is health? Whose views should we take measure and value when constructing quality of life (QoL) measures? Such an

24

enterprise is inevitably a multidisciplinary activity, but economists have played a considerable role in developing both QoL instruments and global measures such as Quality Adjusted Life Years (QALYs).

The demand for health care (box C) is affected by the determinants of health (e.g. income status and work activity) and by the nature of health itself (i.e. by boxes A+B). These factors influence the nature of the demand for health care by different groups (poor/rich, old/young, male/female, majority/ethnic minority). There are many barriers to health care consumption ranging from prices at the point of consumption (i.e. user charges) to time costs and psychological obstacles to care. The agency relationship (or the principal–agent relationship) influences demand because, as a result of the asymmetry of information, patients and carers delegate the task of making demand decisions to 'experts', doctors and other health care personnel.

The subject matter of box D, the supply of health care, is broad, encompassing the markets for factors of production (labour and capital), the health care production function as well as issues of skill mix or labour substitution. The inadequacies, in terms of quantity and quality, of studies into the scope for substituting nurse inputs for doctor inputs is as remarkable as the observable shifts in practitioners' roles in most health care sectors. Perhaps the most important element in this box is the subject of incentives. Incentives can induce behaviours consistent or at variance with policymakers' objectives. They may induce 'on the job leisure' or wasteful and inappropriate activity amongst doctors and health care teams. They may, or may not, induce the adoption of evidence-based practice. Such effects are of central importance in any health care system but are often ignored. Furthermore, the impact of incentives other than doctors and other labour inputs (e.g. nurses) is of great importance. The markets for pharmaceuticals, medical equipment and capital are highly regulated but poorly evaluated.

Microeconomic evaluation at the treatment level (box E) is an activity that preoccupies many health economists. This involves the economic evaluation of alternative ways of delivering care (e.g. choice of mode, place, timing and amount at all phases of the health care process, detection, diagnosis, treatment and after care). Provider industries (pharmaceuticals and medical devices) are a major funder of such activities. Its development, as an area of health economic activity, is a product of provider industry support and of the increasing recognition by clinicians and policymakers that cost-effectiveness is an unavoidable aspect of decision-making when the policy objective is maximization of population health from public health budgets. Unfortunately, clinicians often think that economic evaluation is synonymous with health economics!

If efficiency is the policy objective, the results of economic evaluation inform the choices of suppliers and purchasers and the creation of market equilibrium (box F). Balance between purchaser needs and supplier capacity is reached by a mix of financial prices, time prices (i.e. waiting for GP consultations and outpatient and inpatient treatment) and local practice norms (i.e. non-price rationing). These systems have direct (e.g. prescribing charges reduce demand) and indirect effects (e.g. lack of ready access to primary care 'out of hours', increasing the demand for hospital accident and emergency care).

The evaluation of the entire health system (box G) infers the identification of system objectives (e.g. macroeconomic cost-containment, efficiency and equity, however ranked) and the analysis of the performance of the 'outcomes' of boxes E and F on them. This may involve national or international comparisons and is fraught with difficulties because of problems associated with definitions and the quantity and quality of data.

The final box (H) involves the evaluation of mechanisms to change the health system, in particular how budgetary controls, workplace norms and numbers, regulation and incentives affect performance and how they interact. Evaluation of the efficiency of

26

planning, budgetary and monitoring mechanisms requires either explicit policy goals or assumptions about their identity and ranking.

It should be noted that the terrain of health economics is constantly evolving given the intertwined nature of the global economy characterized by globalization as well as the increased use of Information Communication Technology (ICT). Climate change as well as its implications for global public health is a topical issue that increasingly finds its way into contemporary discussions of health economics. An interdisciplinary approach to health economics research is instructive given the failure of policy prescriptions from specific disciplines.

Further Reading

CADTH Lecture-An Introduction to Health Economics by Scott Klarenbach, MD, available at: https://www.youtube.com/watch?v=Br6bxQW7xBU

The Economics of Healthcare: Crash Course Econ #29, available at: https://www.youtube.com/watch?v=cbBKoyjFLUY

Jakovljevic, M. and S. Ogura (2016) Health economics at the crossroads of centuries – From the past to the future, *Frontiers in Public Health, Vol. 4, Article 115, pp.1-5*, available at: https://www.frontiersin.org/articles/10.3389/fpubh.2016.00115/full

Kernick, D. P. (2003) Introduction to health economics for the medical practitioner, *Postgraduate Medical Journal, Vol. 79, pp.147–150*.

Maynard, A. and P. Kanavos (2000) Health economics: an evolving paradigm, *Health Economics, Vol. 9, pp. 183–190*.

World Health Organization (2003) Learning to Live with Health Economics, available at: https://apps.who.int/iris/bitstream/handle/10665/107690/E88172.pdf;jsessionid=3BD8CF343A6C65A9FACA8E7C0551C20F?sequence=1

3   Demand and Supply of Health

Learning Objectives

By the end of this chapter, the learners should be able to able to:

a)   Distinguish between distinguish between health and health care as they relate to demand

b)   Analyze factors that affects the demand for healthcare

c)   Use Grossman model to explain how demand for health care is attained

d)   Provide empirical evidence on the factors affecting demand for health care

e)   Apply the principle of elasticity to guide in various health care policies adoption.

f)   Demonstrate how supply of health care is arrived at

g)   Evaluate factors that affect the supply of health care

h)   Explain supply of health care using Physician induced demand model

i)   Analyze production of health and health care and relate it to medical technology

j)   Appreciate the role of costs in healthcare production and apply it in the in the perspective of hospitals setting and efficiency.

What is the purpose of demand analysis in health and for health care?

The main purpose of demand analysis in both health and health care is determine on average those factors affecting person's utilization of healthcare service in with the aim improving health. Demand analysis in healthcare fall mainly in the medical services though other dimensions that affects the health of the population is of interest. The

28

reason is at any one given time it is assumed that the choice of medical care is affected by many factors. Hence the health seeking behavior can only be understood from the perspective of economic, social and environmental factors including the provider characteristic.

What determines the patient demand for health care?

The demand for health care is a derived demand. Individual demand healthcare in order to improve their health or health of family members or community. Demand for healthcare is a signal of demand for good health status. Health is defined as a physical and emotional wellbeing of an individual or a defined population. It has a connotation of absence of diseases and good mental health. On the other side Healthcare refers to good and inputs that are used to produce health. Hence, the utility function of healthcare is an indirect one. Thus the patient demand for health will first and fore most depend on two key issues-

a)   The incidence of illness or need for care, this will determine whether individual seek care or not. This is the first stage of demand for healthcare commodities in a health care markets.

b)   The role of the physician as advisor to the patient and as a service provider. Once the individual seeks for medical care, he or she faces the demand for healthcare that will be determined by the physician.

Hence, empirical studies on demand for medical services should demonstrate two things:

1.   How different factors affects the patient's demand for medical care

2.   What determine how the physician will provide care for a given treatment

Model of demand for health and health care

The demand for health is first and foremost analyzed in terms of utility analysis that is based on consumer demand for goods and services.

Traditional theory of consumer demand assumes that consumers purchase goods and services to maximize their utility that is

$U = f(X_1, \ldots X_n)$

Hence, demand for goods X_i is as a result of maximizing individual utility. Assuming that one of the inputs of the utility function is health good, it implies that factor affecting health goods will feature in the unitary model indirectly. What this model suggest is that if the commodity demanded by the consumer is good health, then health can be produced by

a)   Goods and services purchased in the market

b)   Time devoted to preventive measures.

This framework of demand for health and health care was elaborated by **Michael Grossman (1972).** Let us now turn to a model that explains why individual demand health care or medical services.

Grossman Demand Model:

Michael Grossman (1972) was interested in knowing why individuals demand for health care. According to him, demand for health care comes out as a by-product. People demand inputs to get good health. Thus, good health was considered a fundamental good. In essence, health is produced by combining *market goods* and *health time*

Grossman considered the stock of health inputs as health capital. He analyzed the demand for health as a fundamental commodity in an intertemporal frame work, where stock of health was for more than one period.

$U = f(H_o \ldots H_n, Z_0 \ldots Z_n)$

Where;

U is the utility to be maximized

Ho is the inherited stock,

Hn is the stock of health in the ith period,

Z0 is the initial consumption of other goods,

Zn is the consumption of other goods in the ith period

The model enables the determination of optimal stock of health that is required by the consumers using intertemporal choice model. The model is based on the assumption that:

a)   Individual has information about the value of their health stock

b)   They are assumed to know the rate of depreciation of their health stock

c)   They know the health production function that they face i.e., inputs and output at their disposal are known.

Hence, the inputs that goes into health production are the key drivers in the demand for health care. It is the nature of these assumptions that are the basis of critics of Grossman model of 1972.

In his analysis he was interested in asking, why do consumers demand health?

To answer this question, he provided two reasons:

1.   As a consumption good health is consumed directly which mean people who are healthy have taste and preference and they enjoy consuming good and services which make then happy. Hence people are happier when they are healthier because the also enjoy leisure time.

2.   As an investment he claimed that people demand so that they can have health days' work. Good health permits people to do other things that produces return including earning income which they can invest to produce more health days and enjoy life. He used human capital theory which says that individual invest in themselves through education, training and health to increase their earnings.

Figure 3.1: Investing in health Capital

- Health care
- Diet
- Exercise
- Income
- time

Health inputs

Health capital stock over time

- Length of years with good and bad health

Health outputs

- Healthy days
- Physical health
- Activity limitation
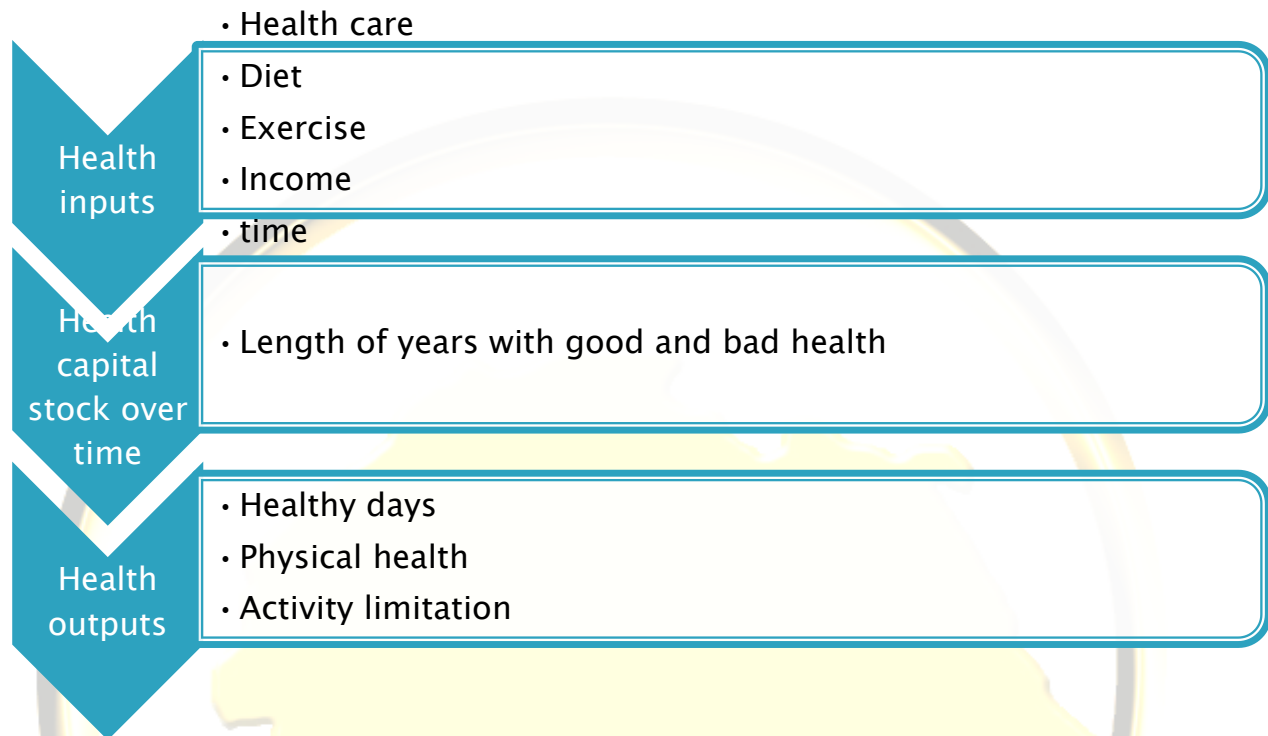
The schematic diagram explains the concept of health capital. The diagram shows that the consumer applies health inputs as investment in health capital. These investments improve or maintain consumers' stocks of health. The second low explains that overtime the health stock s may grow, remain constant, or decline, either slowly with age, or more precipitously with illness or injury. While the third row explains the goal of a consumer as health outputs governed by healthy days. This goal of healthy days dictates how much time and money to invest in health stock.

What does this model say?

a)   When people get old, their stock of health depreciates at a faster rate, so investment must occur at a faster rate; i.e., the old spend more on health care than the young.

b) Higher wages imply that time spent in the unhealthy state is more costly. Additionally, higher incomes increase the consumption value of health. Health spending rises with wage and income.

c) It is hypothesized that more educated people are more efficient at producing health. Spending on health care, then, will fall with education.

Grossman findings can be summarized as in the following:

d) People want health; they demand medical care as an input to produce health.

e) Consumer produces health by combining time devoted to health improving efforts with purchased medical inputs

f) Health lasts for more than one period. It does not depreciate instantly

g) Health can be treated both as a consumption good and an investment good. As consumption good it makes people feel better and as investment good it increases numbers of healthy days available for work and to earn income.

Other comparable approaches to demand to determinant of care health care demand

Fuchs (1982) in his paper introduced time preference and health into relationship between intertemporal choice and health status. Intertemporal analysis assumes that individuals incurs cost today to get benefits in future hence implying that they are investing. According to Fuches people differs in their preference to invest in health. Some want to consume today and others saves to consume in future. Such difference can explain variations in health in two ways.

a) People with low time preference: These are people with high very appetite for current consuming. They believe tomorrow to be very uncertain or random and they would rather deal with current consumption instead of wait and see attitude

b) People with high time preference: The assumption is that people with high time preference would invest in many years of schooling. Given that education increases health enhancing activities. He found a positive relationship between Time preference and schooling

Another intertemporal finding was done by Taubman and Rosen (1982) on the determinant of health status. They used cohort study and found that Single men are likely to be in worse health compared to married men. Single men are likely to be in better health state than divorced men. Married men dies less than single men. According to their findings spouse education had positive influence on health (elementary level but not at secondary and above). Hence these factor are critical in assessing the health production thus because they serve as critical inputs that individual knowingly or otherwise demand to influence the demand good health.

Other Key determinant affecting demand for health that need a clear discussion and empirical re-assessments?

Many empirical research has provided estimates of the relationship between patients and physicians factors and use of medical care. These factors have been addressed in the literature showing that the demand for health care as a function of individual, provider's characters and institutional arrangement which is proxied by government policy. Individual characteristics include variables such as:-Income level of the individuals or the household, occupation of the individual, education level, age, sex, marital status, household size, cultural values including religion, peer effect including network of the individual, severity of illness, Trust on the provider and , ownership of health insurance.

Provider characteristics that has empirically been found to affect demand for healthcare includes but not limited to:- Distance to the provider, provider price, quality of provider services including process(how one is treated by the nurses for instance), physician training, education and specialty, reputation of the provider, availability of provider substitute(for example, existence of traditional herbalist or traditional midwifery for maternal health) and location factors. Institutional or government policy includes implementing health related policy that could have effect on either demand or supply side of the healthcare market. Government policy affects either consumer or the provider or both.

Note: Empirical findings will be shared and discussed in class.

Assignment

a)   The frequency of health care visits are often used to measure service demand but there is a lot of criticism in the use of this variable. Discuss the strength and weakness of using visit as a measure of healthcare service demand.

b)   Explain why a high demand price elasticity in healthcare market implies that the market is competitive and is always true.

c)   Discuss the Grossman model of demand for health care in terms of its applicability in consumption and investment of individuals in an economy.

d)   Using study undertaken in your country evaluate factors that have been found to affect the demand for healthcare.

Supply of health care

What is supply? Supply is usually defined as the amount of goods and services firms wish to sell at alternative prices. In a competitive market, the point at which demand and supply are equal indicates both the price at which goods and services will be exchanged, and the amount exchanged. Supply plays a subsidiary role to demand in most economic application. For example if demand for a product changes because of say peoples tastes changes, firm will produce less; again, if people demand more of a good, firms will increase production and /more firms will enter the markets to meet this additional demand.

In traditional economic model, demand is key while the role of supply is rather passive. If for example, demand shifts outwards, causing a short-run increase in prices, quantities and profits, supply will increase to meet this demand. In contrast, a change in supply for example due to a reduction in the cost of production does not influence

people's demand curve. If policymakers want to change a competitive market for some reasons, traditional economics would suggest focusing on demand. The main tools would be a) **Price** maybe by the way of taxes to suppress demand and b) **improving consumers' information** so that they can make choices that are in their best interest. Beyond that the list of demand-side is rather thin.

Does this principle always apply in supply of health services?
Not really. Why?
Health services market differs from other commodity markets in the following respect, supply considerations play a very important role in determining both prices and output. The suppliers are not just passive producers of whatever is demanded. From example in the case of the physician who acts as the patients agent, suppliers can play a major role in determining what services the consumer ultimately chooses to purchase.

Unlike in the case of traditional economics, suppliers in the case of health services are not passive actors in the operation of the market, the list of policy tools available is considerably expanded. Among the most prominent set of these in the health services area are incentives designed to influence providers' behavior especially using different payment mechanism like fee-for-service, capitation, salaries etc. Unlike in the traditional economic model where supply is quite passive, supply factor in the health service market signal demand.

What does supply of health care entail?
The supply health care is understood from the dimension of Production function. Economists often talk of output being produced using a production function that uses labor, capital, and intermediate inputs. The production function describes the technical relation between the outputs resulting from inputs. The amount of health care produced is constitute the supply. This supply of health care services will depend on the following factors that will be discussed in details in the production of healthcare section;
•   Doctors

- Other health care workers, e.g. nurses, dental hygienists, administrative staff, technicians, staff in care facilities such as hospices and nursing homes, and many others

- Health care capita, i.e. establishments that provide health services e.g. hospital doctors' offices, clinics, nursing homes, and so on.

- Technological progress: Technological progress comprises both product and process innovation. By product innovations, we mean increases in the types of goods and services available to households and doctors.

- Vast array of drugs now available on the market, which is the outgrowth both of research and development at pharmaceutical companies and of publicly funded research

- Advanced machinery used in modern health-care facilities. For example, surgeons today can perform operations that surgeons of previous generations could not.

- Government regulations. Example Licensing which provides more than a guarantee of quality but also limits entry into the profession

Theoretical assumption of dependency between demand and supply in the market for health care or what literature refers to as supply-induced-demand

Under the traditional competitive market model, supply and demand are assumed to be independently determined. This case is relaxed in the health care. There is a perceived relationship between demand and supply of health care services through suppliers and more so the physicians. In medical care demand the patient selects a physician who then makes choices on medical needs for the patient. This decision is based on physician's knowledge of the patient's financial resources and medical need. With the growth of more comprehensive insurance policy, financial constraint has remained to be less important, the physician is thus operating as the advisor to the patient as well as a supplier of the services. The focus now is on whether suppliers and particularly physicians act totally in the patient's interest, or if alternatively they can succeed in convincing their patients to act a way that also benefits them.

This introduces the Demand creation hypothesis which state supply of physicians in area cause price and quantity of medical care to increase through the demand inducement or Supply Induced Demand (SID) which simply means change in demand associated with the discretionary influence of providers over their patients. This is contrary to the traditional economic theory which state that increase in supply would lead to a decrease in price and an increase in quantity demanded for a normal good.

Demand inducement refers to a "physician's alleged ability to shift patients' demand for medical care at a given price, that is, to convince patients to increase their use of medical care without lowering the price charged. Hence this supplier induced demand (SID) can be defined as the amount of demand that exists beyond what would have occurred in a market in which patients are fully informed. In the health care market **Provider-induced demand** (PID) is an economic term that refers to a greater **demand** for services than what would otherwise be expected in a perfect market.

This has been questioned in terms whether physicians act as perfect agents for their patients, or whether they can actually induce demand among patients for their services.

What are the causes of PID?
PID in the health care market can occurs due to several reasons;
Firstly, due to asymmetric information which means patients lack knowledge on amount of medical care they need due to illness. The consumer is not able to judge the quality of service because the information is costly to acquire. It requires investment in medical education which takes a longer time than other professional (training as a doctor is expensive). The existence of SID needs an agent to work on behave of the consumer to ensure that the consumer maximizes utility. In medical field, it exists because the agent maximize their own utility at the expense of the consumer due inability of the consumer knowledge on what they need to satisfy their utility.

Secondly, learning by experience is limited. A consumer cannot learn by experience for frequency of episode is rare. Only people with chronic illness can learn from experience. Hence, they could be having an idea of what they actually want over time in terms of medication though not when the condition become severe. Hence the existence of this limitation lead to need for an agent relationship.

Thirdly, when there is a falling revenue from the provider. This could stem from increased competition where providers have to promote more procedures to patients. Given a health care market if all the provider behaved the same in terms of using inducement, there will be an increase in aggregate quantity of care that does not correspond to healthcare need. In such a case on need to note that PID will take a shift in demand instead of a movement in demand curve.

In a simple market supply and demand increases in response to increased competition. E.g. An increase in supply from say S1 to s2 implies an increased quantity consumed from say h1 to h2 in Figure 3.2. This is implies a shift in supply curve that relates to more services supplied under the same price.
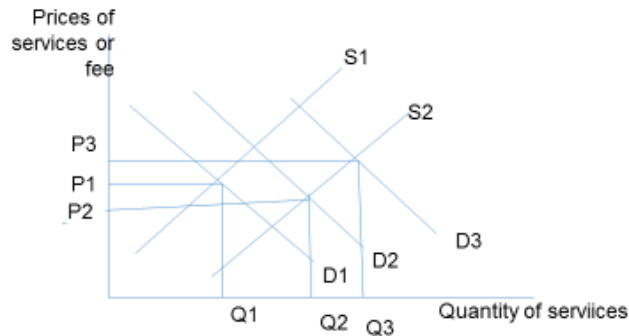
Figure 3.2: Supply and Demand
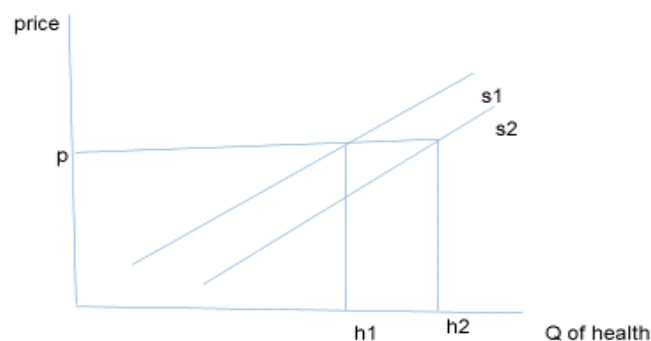
But according to Reinhardt (1978), this cannot auger well with a provider who is a profit maximizer as demonstrated by

**Figure 3.3: Supply and demand model of Physician Induced Demand**

According to Reinhardt, if all physicians in the market behaved like our individual physicians, an increase in the supply of physicians would lead to a increased quantity of care. But for him he assumed a situation of forces of demand and supply behaving as in tradition economic theory where increased supply would mean a decrease in price. Hence an increase supply from S1 to S2 implies an increased quantity consumed from Q1 to Q2 and a decreased price from P1 to P2. Thus, the only way that a provider can rise price to a level higher than before the supply increased could only be caused by inducement. This then implies that both price and demand will increase to D3 and P3 respectively. Thus, in the words of Reinhardt SID is identified when demand increases sufficiently to cause the price (fee) to rise. This argument was supported by Feldman and Sloan (1988) who attest that quality could be the key incentive which is also costly to implement.

The conclusion based on the assumption that physicians attempt to maximize their income can be disputed on the account that other things could also matter for the physician for example availability of leisure time and professional ethics(having the patient's interest at heart or alternatively , distaste towards inducing services).

If a physician tries to maximize both income and leisure jointly, his or her supply curve may exhibit a "backward –bending" shape. At low levels of fees, for example, point A in Figure 3.4 higher payment rates will induce more work. At some level, such as point B, unit payment rates are so high that the physician may decide to enjoy more leisure. Thus, a reduction in payment rates would lead to an increase in the quantity supplied (supply from the physician side). Note that this is the same results as under demand inducement- but without there being any inducement whatsoever.

Figure 3.4 Backward Bending Supply curve
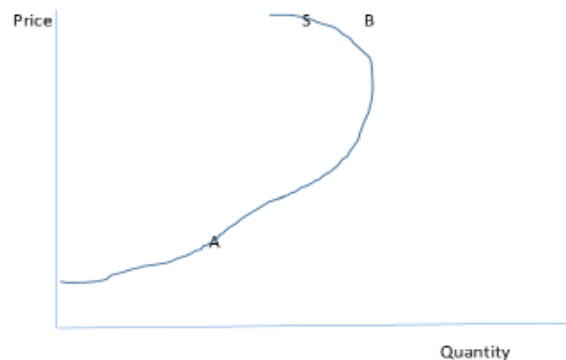
Factors that affect the size of SID in the market for health care

1.   **Method of Provider payment or reimbursement**: The way in which a doctor is paid is a key driver to the size of PID. Let us look at two reimbursement mechanism;

a.   Fee-for-service

b.   Prospective payment system

Fee-for-Service. In FFS services are paid for in terms of quantity or amount provided. Here

Total Revenue=Price *quantity. Total revenue increase if either price or quantity or both increases. If we assume that Price is constant, then quantity must increase to ensure

42

that that revenue is unchanged. Hence under this scenario we can talk of SID/ PID by seeing what happens to quantity. In this system there is a positive relationship that SID will be present for doctors know that the more the patients the more their expenditure which in turn becomes doctor's income. Hence there is an incentive in the side of the doctor to induce demand.

Prospective payment system: Under this system, the patients or health authority enters into contract with a provider. To provide defined set of health care services to a population group for an agreed amount of money. This amount is fixed in advance. Very common in America and Europe. The provider has a duty to provide those services for that budget. Hence, they have an incentive to be efficient. Most of the providers especially doctors are paid in terms of salary. This is invariant with the amount of services that the doctors provides, hence the quantity incentive is absent. The implication is that the PID is highly unlikely under prospective payment method.

2.   **Nature of the supply system**. If there is if excess supply of doctors and finite population of patients the possibility of having SID will increase. This is due the fact that doctors will have to compete for the finite market. The average income for doctor's decreases population of patient is assumed to be finite. The only way to raise more income is to increase services per patient. In a way they have to increase the quantity for increasing the price under a competitive market will decrease the number of patients visiting them while holding other factors constant. Hence, they have to use some form of inducements to increase demand for their services which lead to PID.

3.   **Cost to doctor or cost inducement of PID**. In most health care system inducement is costly to doctors. Doctors are supposed to adhere to some ethics as read when taking their oath of work Thus if found and proved practicing PID the cost to high in some instances it leads to doctor deregistration from performing professional work. If the cost of inducement is high the possibility of PID is low compared to when it is that cost is low.

4. **Incentives faced by the consumers**. If the consumer has to bear the full cost of investment in health, then the inducement tends to go down. This is especially so with the case of out-of–pocket payment where the consumer has no insurance. This is because even if the consumers are not informed, they are likely to complain on costly alternatives which will act against PID. Alternatively, where consumer is 100% insured, he/she does not bother of PID. But in the case of co-payments (this is where the patient does not pay 100% nor the insurance company for the treatment of the patient) it will depend on the contribution of the patient to an insurance policy. If for example the patient pays 40% and insurance pay 60% the PID will depend on the income of the patient. If insurance company finds that simple ailment has become prevalent, it will exclude payment of the trivial though this will be clearly stated in the insurance policy.

5. **Frequency of purchase**. This implies the occurrence of illness among the individuals in the population and readiness to get treated. If the occurrence is infrequent like in surgery cases, then the incentives for agent to induce demand are high. In the infrequent cases one doctor could be getting one operation in a week or in a month. Agent are thus likely to initiate a surgery as much as it might not be the right time to do it or is not necessary. A good example given is the caesarian cases where pregnant mothers are recommended for operation without actual need. On the other hand, PID is said to have a low probability case which are frequent like with chronic illness mostly because the consumer has some level of knowledge on what they need due to their past and continuous learning.

6. **Presence or absence of Peer groups reviews**

Presence of peer reviews ensures that some standards are set mostly in terms of utilization or certain medical procedures. This peer review relies on the second opinion. This peer review is mainly concentrated among the insurance lead payments. The peer review relies on the second opinion. In situation where peer review on utilization is absent on expect a rise in SID (good example is in the case of a dentist). Where peer review high the possibility of SID is very low if not non-existence.

44

Supply of hospital bed hypothesis

"A build bed is a filled Bed" hypothesis was propounded by Roemer and Shaina(1961).Building beds result to them being filled with patients whose hospitalization is not medically indicated. Hence policy wise the number of beds in a area should be based upon a medical determination of the need for beds. Medical need assumed to be same in all region hence beds should be based on population. Supply- create -demand hypothesis and its policy implication has conceptual and empirical problems. In part a pure correlation between beds and hospitalization could imply that they were both affected by same factors. It is a fact that areas with high demand for hospital care generated an increase in supply to satisfy that demand. Areas with high demand would have a larger numbers of beds per capita than areas with low demand. Some possible theoretical explanation for this correlation are:-

- A current demand for hospital care may exceed supply in which case increasing the supply of beds satisfies existing demand.

- Again once the bed has been filled more bed will be built in anticipation of higher demand.

Assignment:

a) In many health economic textbooks and empirical work on supply for healthcare, supply induced demand and Physician induced demand are used interchangeably. Clarify if the there is any difference between the two.

b) Assuming that SID is prevalent and substantial, what are the implications for policy? Policy to reduce national insurance? Policy to increase the number of medical schools graduates?

c) If physician fee is fixed so that they do not adjust to changes in supply, explain how SID can be manifested.

d) What forces limit a provider's ability and willingness to engage in SID?

Elasticities and healthcare good and services

The objective of this sub-section is help our demonstrate the use of elasticity at explaining the relationship between goods and services in healthcare setting that arises due to changes in the health care variables like price and income. The section also demonstrated different types of health care goods as defined through elasticities effect.

What is elasticity? (A revisit);

Elasticity is defined as the percentage change in the dependent variable resulting from a 1 percentage change in the independent variable under study. In health economics we are particularly interested in the responsiveness of the quantity of care demanded when price of health care changes. Our quantity of care is our dependent variable while price is our independent variable. Hence economists uses the term elasticity to describe the responsiveness of any term y(e.g. quantity of health care demanded) to change in some other variables x(e.g. price of health care).

Thus one can define elasticity as amount of change that occurs in economic behavior when one factor is changed or simply as responsiveness of quantity either demanded or supplied due to a %change in a determining factor

Let's looks at several types of Elasticities that are relevant to our health care setting:

- The Price Elasticity of Demand
- Cross Elasticity and Income Elasticity
- The Price Elasticity of Supply

Price elasticity of demand

A measure of the extent to which the quantity demanded of a good changes when the price of the good changes. To determine the price elasticity of demand, we compare the percentage change in the quantity demanded with the percentage change in price.

Hence in the case of the price elasticity of demand, it can be shown as

E_p= *(% change in quantity demanded)* ÷*(% change in the price)*

OR

$$E_p = \frac{\Delta Q/Q}{\Delta P/P} = \frac{\Delta Q}{\Delta P}\left[\frac{P}{Q}\right]$$

Where Δ refers to change in the variable p and q are price of own good and quantity demanded respectively while E_p is the price elasticity of demand. Note the price elasticity is always algebraically negative because an increase in price leads to a decrease in quantity demanded. When the price rises, the quantity demanded decreases along the demand curve. Price and quantity always change in opposite directions. So to compare the percentage change in the price and the percentage change in the quantity demanded, we ignore the minus sign and use the absolute values. Hence numerical values for price elasticity are often reported in absolute values, eliminating the minus sign.
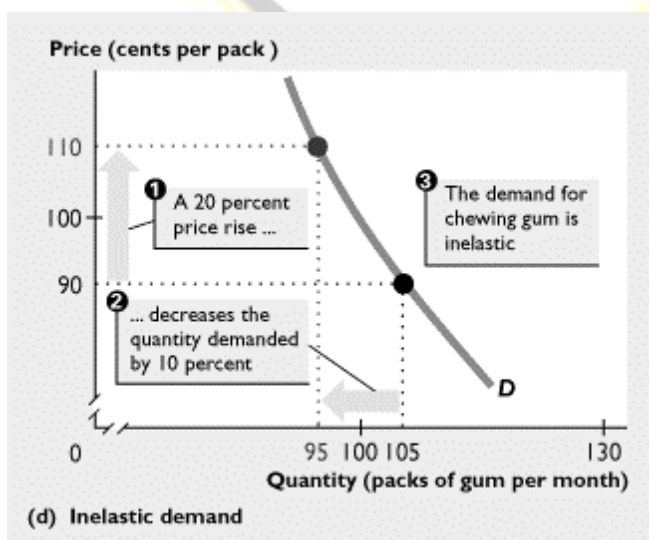
Elastic Demand

47

Absolute values for price elasticities indicates the responsiveness of demand to price in that the greater the elasticity is the greater the responsiveness is. Absolute values greater than 1 and less than infinity are considered elastic. This mean demand is elastic if the percentage change in the quantity demanded exceeds the percentage change in price.
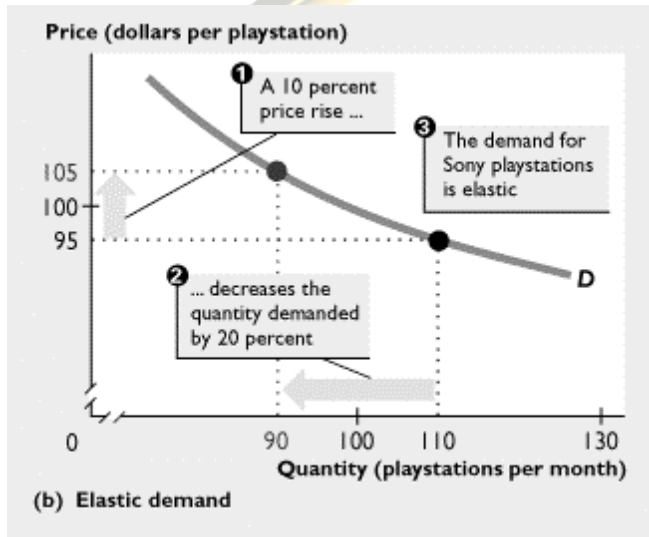
(b) Elastic demand

Fairly inelastic demand

Absolute values between zero and one are considered fairly inelastic. This implies that the quantity demanded remains changes only slightly with a when price changes significantly as shown.

(d) Inelastic demand

Perfectly inelasticity

Absolute values equal to zero are considered perfectly inelasticity meaning quantity demanded does not change with price, here the demand curve is vertical. It implies that the quantity demanded remains constant as the price changes as shown

(e) Perfectly inelastic demand
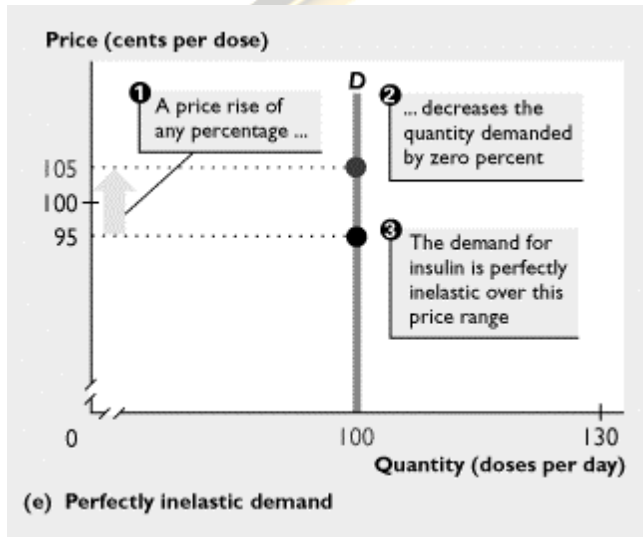
Unit elastic demand

When absolute value equal to 1 elasticity is considered as unity. This implies that the percentage change in the quantity demanded equals the percentage change in price as shown

(c) Unit elastic demand
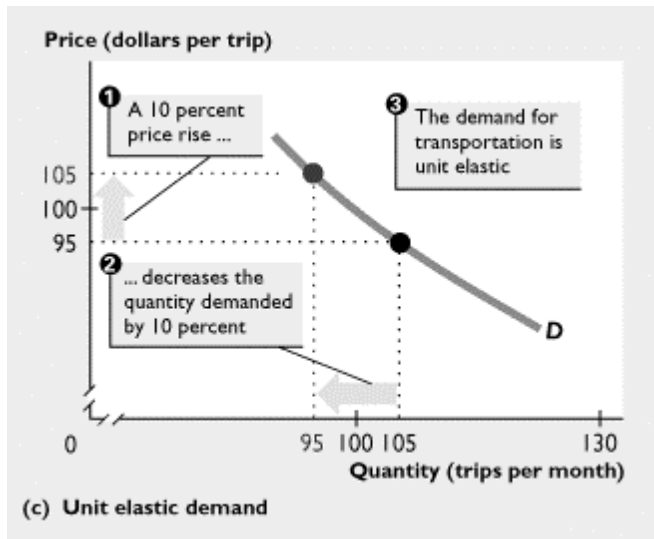
Perfectly elastic

In another limiting case where the absolute value is infinite, elasticity is considered as perfectly elastic and the curve drawn horizontally. This is case when the quantity demanded changes by a very large percentage in response to an almost zero percentage change in price as shown in the figure below

(a) Perfectly elastic demand
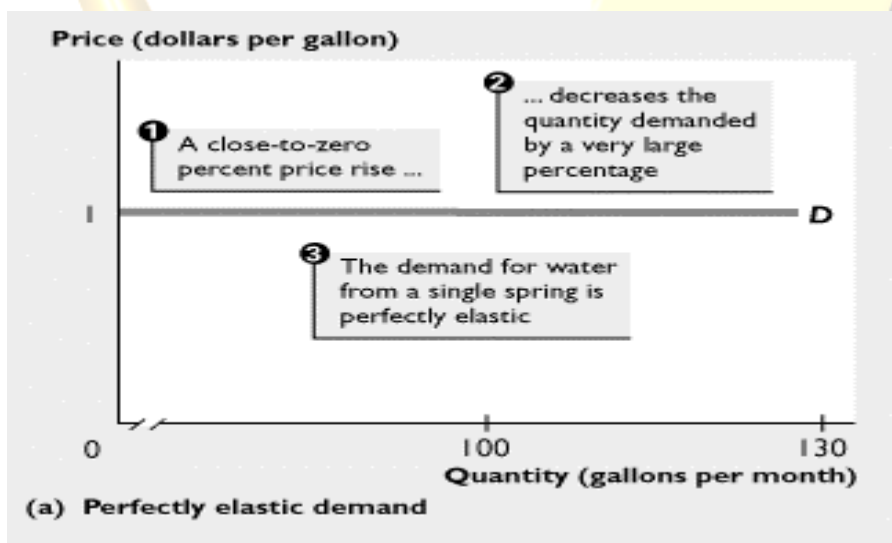
NB: Elasticity can lie between perfectly elastic and perfectly inelastic

Influences on the Price Elasticity of Demand:

Influences on the price elasticity of demand fall into:

• Substitution effects

• Income effects

Substitution Effects

The demand for a good is elastic if a substitute for it is easy to find. On the other hand the demand for a good is inelastic if a substitute for it is hard to find. This explains why scarcity is key in setting prices three main factors influence the ability to find a substitute for a good:

a) Luxury versus Necessity. A necessity has poor substitutes, so the demand for a necessity is inelastic example of a necessity is food. A luxury has many substitutes, so the demand for a luxury is elastic example of luxury good is type of watches or exotic vacations

b) Narrowness of Definition. How the good is defined will determine its elasticity. The demand for a narrowly defined good is elastic example type of food or drinks. On The other hard good that are broadly defined has an inelastic for example food to sustain life.

c) Time Elapsed Since Price Changed. The longer the time elapsed since the price change, the more elastic is the demand for the good.

Income Effects

The greater the proportion of income spent on a good, the more elastic is the demand for the good. This means that the price will have a significant effect on the amount that will be purchased using the available income.

Question: Is slope of demand curve the same as its Elasticity

Answer: **No; Why?**

Slope measures how the quantity demanded changes when the price changes.

Slope depends on the units of measurement of price and quantity. For example, the slope of the demand curve for mineral water   has the units dollars per litter. Slope

Elasticity along a Linear Demand Curve

Along a linear (straight-line) demand curve, the slope is constant but the elasticity varies. Along a linear demand curve, demand is; unit elastic *at* the midpoint of the curve, elastic *above* the midpoint of the curve and inelastic *below* the midpoint of the curve as shown below
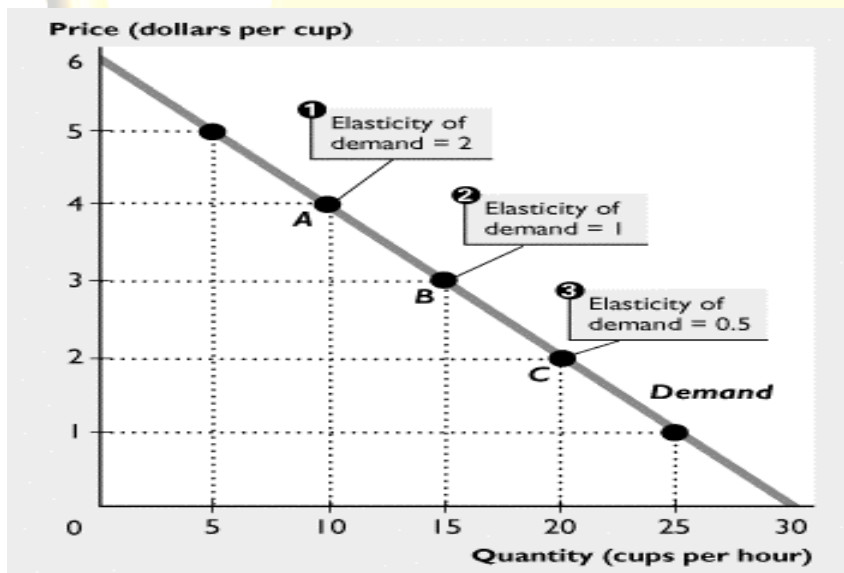
Cross Elasticity of demand

Cross elasticity of demand is measure of the extent to which the demand for a good changes when the price of a substitute or complement changes, other things remaining the same. Thus it relates to relative change in demand due to price change of a substitute or complementary good.

Cross elasticity of demand=%change in Qx /% change in Py

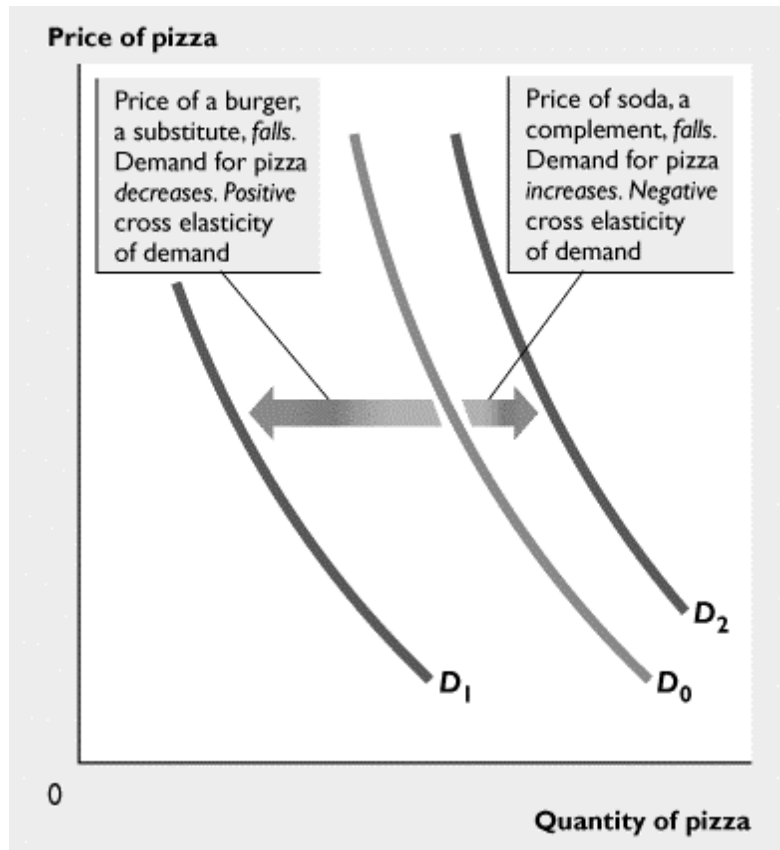Where Qx is the quantity of good x demanded while Py is price of a related good When the value of cross elastic of demand is positive, the two goods in question are said to be substitute (goods that are demanded as an alternative of each other example a *Hedex* and *Panadol , private clinic service vs public health center service*). A fall in the price of a substitute brings a *decrease* in the quantity demanded of the good. Hence the quantity demanded of a good and the price of its substitute change in the *same* direction.

While if the value is negative the two goods are said to be complementary (goods that are demanded less jointly or together, example injecting needle and syringe, an ambulance and a nurses). A fall in the price of a complement brings an *increase* in the quantity demanded of the good. Thus the quantity demanded of a good and the price of one of its complements change in *opposite* directions

Price of pizza

Price of a burger, a substitute, *falls*. Demand for pizza *decreases*. Positive cross elasticity of demand

Price of soda, a complement, *falls*. Demand for pizza *increases*. Negative cross elasticity of demand

D_2

D_1

D_0

0

Quantity of pizza

One need to note that the amount demanded under cross price elasticity changes while the price remain constant. More or less is demanded at the same price.

Income Elasticity of demand

A measure of the extent to which the demand for a good changes when income changes, other things remaining the same. In derivation its shown as,

Ey= (% change in quantity demanded) ÷(% change in income)

This type of elasticity is also called own price elasticity, where the change in demand for the product is as a result of changes in it own price.

$$E_y = \frac{\Delta Q/Q}{\Delta y/y} = \frac{\Delta Q}{\Delta y}\left[\frac{y}{Q}\right]$$

Income elasticity may be positive (in case of a normal good) or negative (in case of an inferior good). If a variable elicit no response after a change in income then elasticity is zero. The figure that follows shows type of income elasticity

54

(a) Income elastic (b) Income inelastic (c) Negative income elasticity
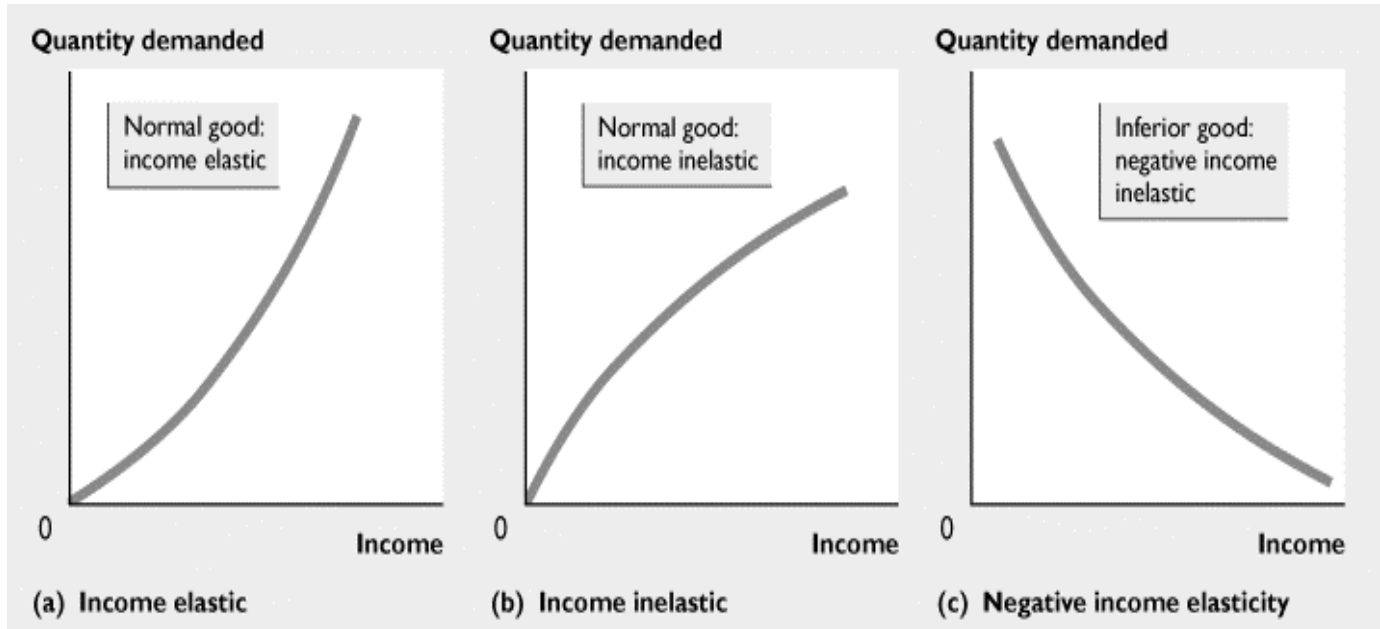
To demonstrate price elasticity and cross elasticity by use a two health care product

commodity	Original quantity	New quantity	Change in demand	Original price(yens)	New price(yens)	Change in price
Colgate	800	1000	200	20	15	5
Close-up	250	300	50	120	100	20

Own Price elasticity of the products

Ep for colgate=200/800*100/5/20*100=25/25=1

Ep for close-up=50/250*100/20/120*100=20/16.3=1.2

A change in price of each respective product will yield a absolute value greater than zero. In case of colgate the elasticity value is considered as unity while for close- up the own price elasticity is considered as elastic because the value is greater than one.

Cross-elasticity on demand for tooth paste

Assume the quantity of close-up changes due to the change in the price of colgate the cross price elasticity will shown as

50/250*100/5/20*100=25/25=1

Interpretation: a positive cross elasticity implies that the two commodities are substitute to each other.

THE PRICE ELASTICITY OF SUPPLY

Just like price elasticity of demand, price elasticity of supply, measure the extent to which the quantity supplied of a good changes when the price of the good changes. To determine the price elasticity of supply, we compare the percentage change in the quantity supplied with the percentage change in price. That is,

E_P= (% change in quantity supplied) ÷(% change in the price)

$$E_p = \frac{\Delta s/s}{\Delta P/P} = \frac{\Delta s}{\Delta P}\left[\frac{P}{s}\right]$$

Again just like price elasticity of demand we also have similar categories of Price elasticity of supply namely;

a. Perfectly elastic supply

An almost zero percentage change in price brings a very large percentage change in the quantity supplied.

(a) Perfectly elastic supply
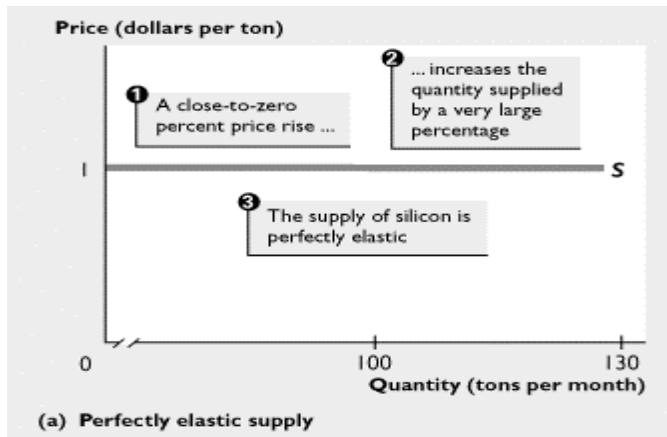
b. Elastic supply

The percentage change in the quantity supplied exceeds the percentage change in price.
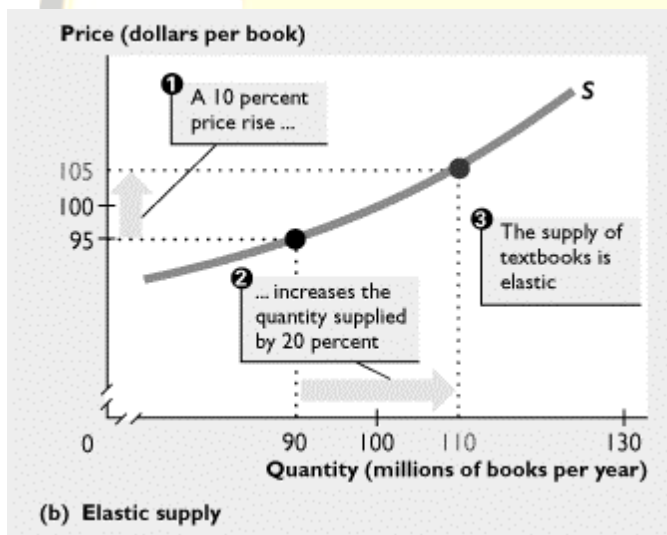
(b) Elastic supply

c. unit elastic.

The percentage change in the quantity supplied equals the percentage change in price.

(c) Unit elastic supply
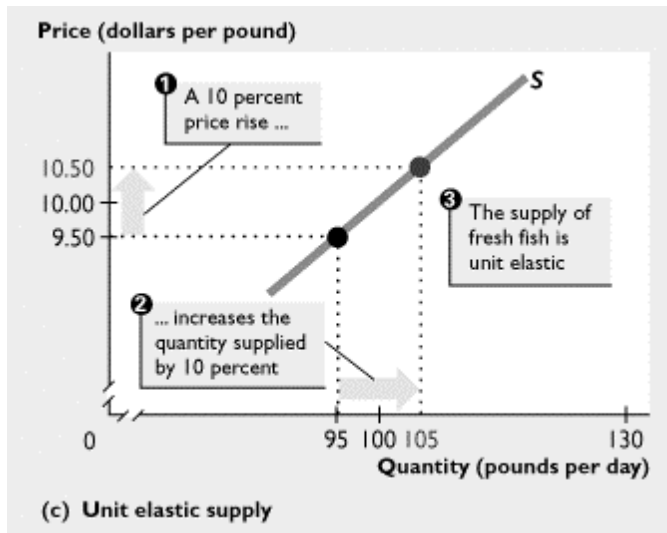
d. Inelastic supply

The percentage change in the quantity supplied is less than the percentage change in price.
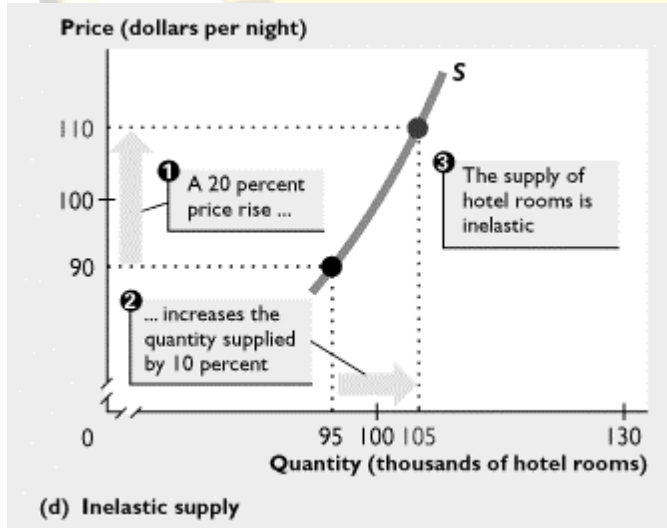
(d) Inelastic supply

e. Perfectly inelastic supply

The percentage change in the quantity supplied is zero when the price changes.

(e) Perfectly inelastic supply
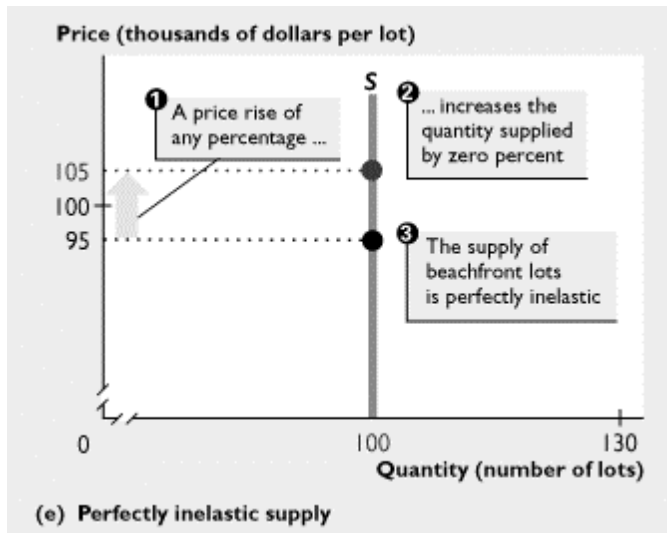
Features of Elasticity of Supply

Factor that influence elasticity	Inelastic	elastic
A rise in price means	A smaller raise in supply	A larger rise in supply
The good is produced	Slowly	Rapidly
The time period	Days	
The firm stock	Limited stocks	larger stocks
Perishability	E.g. Beef	E.g. Screws

Applications of Price Elasticity of Demand in health care promotion

The importance of price elasticity to policy questions can be illustrated with an example regarding cigarettes, which are clearly a health concern. Let assume that the state intends to win two policy objective, one deter cigarettes smoking in order to improve health outcomes of both smokers and passive smokers, two increase tax revenue that can be used to purchase cancer equipment.

In order to achieve this, the policy makers can use the elasticity approach to assess in advance the implication of price increase through increased tax. Using the figure below we notice that the demand for cigarette is somehow perfectly elastic implying(D1) that increase in price will not deter demand for smoking but it will increase tax revenue as

desired. Hence the objective of increasing price in order to promote health lifestyle through reduced smoking cannot be achieved through this policy.

If on the other hand we discover that our demand for cigarette is elastic as shown by D1 and D2(this could a case where youth who are not employed are key consumers of cigarettes) an increase in price through tax rise, will decrease the quantity of cigarettes demanded but again this will have negative effect on tax revenue due to reduced demand. Thus the more elastic the response, the greater the effectiveness of an excise tax in inducing people to reduce their level of smoking. Lewit and Coate(1982) indicates that teenagers are more responsive to cigarette prices than are adults. In such case taxes on cigarettes will be relatively more effective with teenagers.

Price of cigarettes

Tax of 50 cents

Even more Elastic

More elastic

Least Elastic demand

Q3 Q2 Q1

Quantity of cigarettes
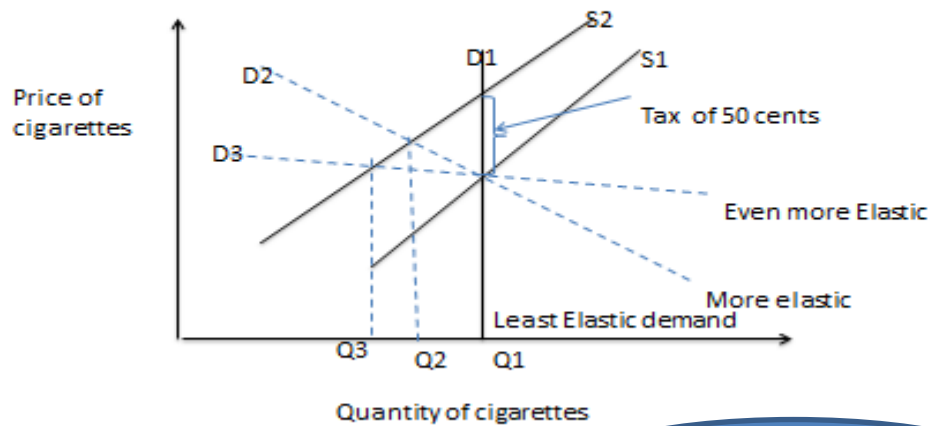
Assignment: Discuss why these scenarios will hold or not hold:

1. Goods we call necessities tend to have elasticities less one in absolute terms
2. Luxury goods have elasticities greater than one.
3. Short-run elasticities tends to have smaller elasticities than long-run
4. Goods that cost a small fraction of individual income tend to have very small elasticities in absolute terms

Some empirical work on elasticities related studies in health care.

Gertler et al. (1987) find that although user fees generate substantial revenue in Peru, they are accompanied by large reductions in aggregate consumer welfare, with the poor shouldering most of this burden.

Sahn et al. ([2003](#)) show that own-price health care elasticities in Tanzania are decreasing in income, suggesting higher fees create larger barriers to care for the poor. Analogously, Deininger and Mpuga ([2005](#)) find that the removal of user fees in Uganda led to improved access to care by the poor and a reduced probability of sickness. Dor et al. ([1987](#)) find that travel time plays an important role in the demand for health care in rural Côte d'Ivoire, with poorer individuals' demand being significantly more travel time elastic than that of the wealthy. All this is consistent with the broader literature on the demand for health care services.

Health and healthcare Production

> **"Health is not everything in life, but without health, life is nothing".**

This is a popular statement that points to why health is important in life hence the need to produce it. This popular statement points to two property of health

1.   Health is a highly valued asset. This is captured by the first part of the statement that is "other thing matter in life but health ranks higher.

2.   Health is a prerequisite for other activities. This means poor health limits capabilities of the affected person both to work and enjoy life

The short coming of this popular saying fail to point what economist notices t that health can be produced thus departing from the traditional belief that good health is a ***gift of God*** and poor health is simply ***bad luck***.

Those who fail to recognize that health can be produced advances two key objections with respect to health production. These objections can be highlighted as;

a.   There is **lack of control** in terms of health production like the way we can control other production processes. In traditional production like agriculture we can control the factors of production or inputs in order to maximize output. For example if a

farmer expect a future increase in demand of their produce, they will produce an inventory, on the other hand, if they expect a drop in demand they will limit production.

But: Process of producing health seems to lack this type of control

b. There is **lack of tradability**- Production of output which is not governed by subsistence motive is usually sold to third party. That would be inconceivable (according to this objection perhaps this can only work in a in a slave economy where prices of slaves were determined by their health)

According to Economist these two objections were found to be weak due to the following argument.

Effort taken to maintain health can be considered as inputs in the health production process. As much as they do not help to avoid illness, they do contribute to better health. Control of health process will depend on invisible capital stock which includes education. Education would for instance cause an increase in wage which is likely to increase purchase of health stock. Again, Individuals lifestyle would have direct effect on the control of health production for instance; smoker vs non-smokers, over-tired people vs rested people, and normal weight vs overweight people, sedentary lifestyle Verse active lifestyle, alcoholic verses non-alcoholic

In terms of tradability, individuals trade with themselves. They trade better health to something else of value. They must sacrifice things such as consumption of tobacco, alcohol or some types of food and time that could have been used for other things in order to maintain good health.

Applying the concept of production function in health points to the fact that different inputs make different contributions to the resulting output. Medical expenditure is one such input whose marginal cost is known to be very high. Economist can only justify high medical expenditure only and only if it is characterized by a very high marginal

productivity. It is important to point here that an improvement of environmental quality could have a greater marginal productivity in terms of health than a continued increase in health care expenditure which then implies just like in the traditional production where government can regulate production through some policies like taxes and subsidies, government can also introduce regulations that can go hand in hand with increasing health outcome of the entire population. A good example is where government can regulate environmental pollution or introduce policies that will affect social determinant of health like education positively. All this will lead to health production.

Production Function of health

Production function is a summary of factor inputs and outputs. Health production function shows the relationship between health care inputs and health outcomes/status. Health status in the production function is assumed to be an increasing function of health care inputs. A possible 'health production function' would look like this;

$$H_i = f(N_i, C_i, I, S_i, S_h, L_i, T_i, A, M_i \ldots)$$

Where:

H_i: health of household member i,

N_i: nutrient intake of household member i,

C_i: consumption of household member i,

I: household income,

S_i: education of household member i,

S_h : education of household head h,

L_i: leisure time of household member i,

T_i: work time of household member i,

A: resource endowment of household,

M_i: health care consumption of household member i.

Question: Does health status depends entirely on health care inputs?

Answer!: No; reason, it also depends upon the population's biological endowments, environment, and lifestyle. It's worth noting that health care includes things like medical

care provided by the doctors of medicine and many other health professionals. Conceptually, the health care (HC) measure may be thought of as an aggregate of all these health care. Marginal contribution of health care is its marginal product.-the increment to health caused by one unit of health care holding all other inputs constant.

Figure 3.5: Health production Function
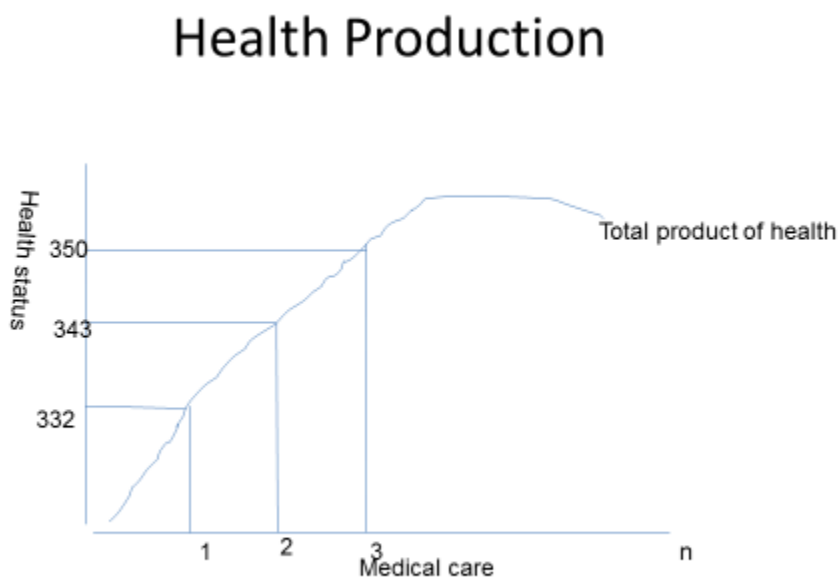
Increasing health from zero to one unit improves health status by say ΔHS_1 e.g. from 332 to 343; $\Delta HS_1 = 11$ health status units. The next unit of medical care delivers a marginal product of ΔHS_2 equal to say 7 etc. These marginals are diminishing in size, illustrating the law of diminishing marginal returns. If the society employs a total of n units of health Care, then the total contribution of health care is the sum of the marginal product of each of the n units. Note the marginal product of nth health care inputs is very small, nearly on the flat curve.

Marginal product of health care

Marginal product of health care is derived by tracing the additional health status brought by the nth health input. Health production function is drawn as a rising curve that

65

flattens out at higher levels of health care but never bends downward. The marginal product of health help to conceptualize if additional health inputs could harm health of the population.

Figure 3.6: Marginal product of health
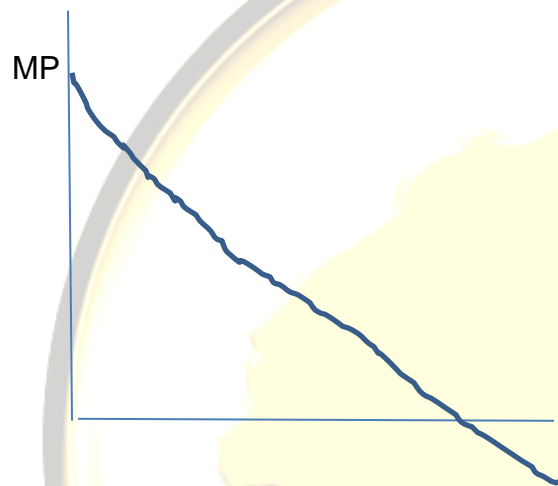
Medical Inputs

This reasoning can be supported by two scenarios

a) Iatrogenic (meaning provider-caused) disease- where too many medical interventions like surgery (where combination of drugs may have unforeseen and adverse interactions). If iatrogenic supersede additional returns from the nth health care the net effect will be a net harm which will be reflected by a negative marginal product

b) Dubos (1960) and Illich (1976) argues that some health care inputs like "medicalization" takes away ability to face the natural hardships of life by "medicalizing" these problems. This lead to less personal effort to preserve health:- the result becomes a decline in the health of the population and thus a negative marginal product for medical care.

It is the marginal that are always relevant to policy propositions not the totals. Example

•   Can ask society to eliminate all the health care spending

- But Reasonable to ask a reduction of health care expenditure by say US$ 20 million and invest those funds to other health uplifting sector like consumption, education, transportation, and defense.

Health care Production process

Production process in health care entails looking at three process that is, inputs, through puts/contents and outcomes. These indicators are helpful in gauging the productivity of our health sector as measured through efficiency.

Table 3.1: Production process of in Health care

Inputs	Through Puts/Content/output	Outcomes
Health expenditure per capita	Average length of stay in hospital	-Under five mobility
-Hospital beds , per 1,000	Bed occupancy rates	-Other mortality rates
Mixed of health personal(Doctors vs nurse ratio)	Bed turnover intervals (time that beds are vacant between patient0	-Morbidity rate
Doctors/ nurses to population ratio	Daily caseload per health staff at different facilities	-effective immunization
Prescribing rate of Drugs	Utilization rates of other facilities	Re-admission rates
Allocation of spending between different services(e.g. capital costs; staffing; supplies Equipment, Lab, Syringes etc.)	Admission rates(percentage of population admitted into a hospital each year)	Patient satisfactory measures
Health	Waiting times for different	Success or cure rate for

Inputs	Through Puts/Content/output	Outcomes
expenditure as a percentage of GDP	procedures	different procedures
Allocation of spending between different services(e.g., MCP/FP, AIDs Care etc.)	Accuracy of diagnostics tests (percentages of false negative or positives)	Post –operative infection rates
Expenditure per admission, per bed-day, per procedure, or per outpatient visit	Coverage of immunization rates for immunization, antenatal care and other preventive work	Pre-operative mortality

If it was an Industrial process we could measure the cost of the inputs (labor, equipment facilities supplies, etc.) then measure the value of outputs (finished products at market price) and then compare the profits achieved by productivity. **But in health care there are a number of complicating issues that differentiate it from traditional industrial production process, for example,**

1. The Inputs are not always homogenous. A patient with TB at one facility is unlikely to have the same severity of illness and response to treatment as patient diagnosed with same disease at another facility.

2. Following the fact that production process (i.e. health care activity) cannot be uniform, it must be allowed to vary according to specific need of the patients.

3. The outcome will be related to a number of factors independent of treatment/prevention process itself, such as the overall level of the health of the patient, the degree of their compliance with treatment, their home, environment, income levels etc.

4. Health care activities have many goals, not all of which are easy to measure for instance, Palliative care (specialized medical care for people living with a serious illness), aims to reduce suffering, but have no effect on length of life. Its impact in terms of reduced suffering will be hard to measure. Other goal could be

   - Reduction of morbidity rate or illness
   - Reduction of mortality
   - Universal coverage of health, etc.

5. Some indicators in the health care are not always straight forward, for example,

(a) Expenditure patterns in facilities if health expenditure is relatively low in hospital A for instance, is that because the health care sector is efficient is just underfunded. If the outcome indicators are comparable with neighboring countries despite lower expenditure is this because the health sector is more effective, or because of other factors relating to income distribution, consumption patterns infrastructure etc. Similarly, are low admission rates good or bad? They might be good if they reflected a well-functioning Primary Health Care (PHC) system or bad if they reflect the lack of confidence which the population has in the hospital system.

(b) Quality of the information: Because more healthcare is not always better, we should ideally connect information about inputs, through puts and outcomes to the need and to the health and welfare gain which they produce. Is the information available capable of addressing this process? To do this we need to

(i) Know how information was collected and establish whether it is accurate.

(ii) Establish what information is missing e.g. how many were admitted and how many were turned away or failed to be referred or were admitted inappropriately.

The collected information should answer the following questions:

(1) Do the service mirror the burden of disease and what is known about the effective ways for preventing or treating it.

(2) Can the information help us make connections between general health indicators (such as mortality and morbidity) and the activities of the health sector?

NB: Taken together, these general indicators will suggest some areas of possible poor performance and this can be the basis for more detailed efficiently studies at lower levels of the system.

Cost function in healthcare

As noted earlier, production function describes the input –output relation while the cost function describes the cost out-put relation. Cost function is derived from production function through assigning prices to the factors inputs used to produce a given level of output. In the production function Isoquant traces the number of visits/ cases that a physician is capable of treating in a specified time period**. In our factor input plane there are infinite number of** possible isoquant as shown in our Figure 3.7.

The lowest and highest isoquants show that there are many combination of inputs (labour and Capital) that are capable of producing say Q0=100, Q2=150, and Q3= 200 physician visits. But the question is what does it cost to produce these visits and what is the least cost that can be incurred to produce such visits? How much it cost to treat the patients is shown by the straight line touching the y and X axis. The line provides the available resources that can be used to produce a certain maximum of visits. This line is our Isocost curve. Hence in order to produce the maximum number of visits, the isocost must be in tangency with the isoquant as shown in our Figure 3.4. Hence, Isocost curve are straight line just like budget lines. They trace combination of capital and labor that together cost a given amount. Thus, it represents total cost. Our isocost line are shown by L1,L2 and L3.

Figure 3.7 Expansion Path
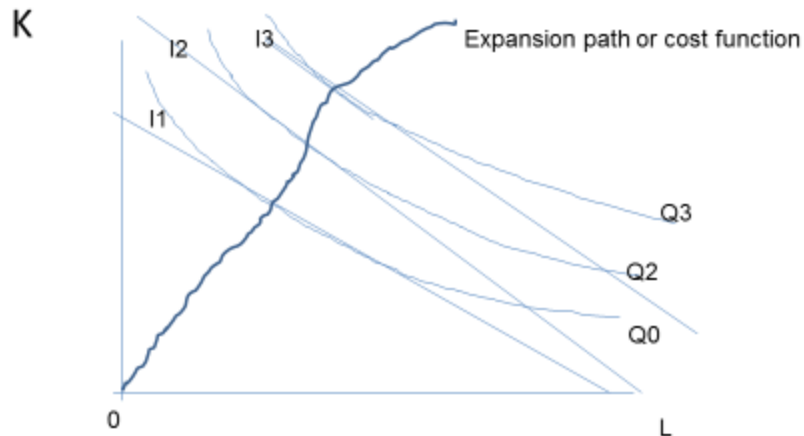
We can define our Isocost in terms of our input as

We start by assuming that Total Cost can buy many combinations of K and L,

TC =rK +wL where r is rental price of capital and w is the wage rate of labor

TC-wl=rK

Which means,

K= (TC-wL)/r this is considered as our isocost line or equation

Back to our Figure 3.7, the firm wishing to produce a given output level/visits say 100 visits will minimize its cost by choosing the lowest isocost curve that is tangent to the 100 visits isoquant curve. Given knowledge of input prices we can calculate the output cost. Assume r=$1200 and w=$1000 and L=20, K=25. Then the total cost for 100 visits should equal to ($12000*25)+($1000*20)= $ 50000. Alternatively, if a firm want to produce 150 visits with 30 units of labour and 40 units of capital using the same input prices the total cost of the visits will be $78000.

Combination of the production function represented by the isoquants and the cost requirements represented by the isocost curves generates a set of outcomes that when

71

joined yield the expansion path. The expansion path provides information that shows us that a given output is associated with a minimum cost. Once these points are represented in a graph they provide us with the firms long run total cost function (see Figure 3.8). The cost curve goes through the zero origin. why? Reason: If a firm decide to produce nothing it incurs no cost. The cost function is a lazy S-shape which is a pattern thought to mimic many firms in practice. Our Total cost function is derived from production function as demonstrated in the figure. This cost function is associated with the long –run period when firms are able to alter their production function in terms of inputs and resources.

The cost function is a "frontier" in that it represents the minimum possible cost of producing a given output. Actual firms may operate somewhat inefficiently and will observe cost levels above the frontiers. This bring the issue of efficiency which will be discussed in a later section.

Figure 3.8: Derived Total cost curve
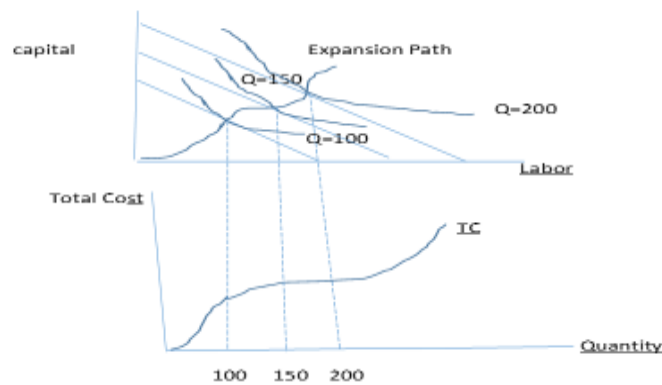
Possibility of factor substitution in healthcare production

In traditional industrial production function economist attest to multiple technique for the production of goods. This implies one can choose for example capital (labour) intensive technique when capital (labour) is relatively cheaper in order to maximize profit. Hence one can substitute one factor inputs to get the same output albeit at low cost.

However, in healthcare production function, practitioners often argue that there is only one right way of administering correct treatment for a given illness. This what Victor Fuchs called "monotechnic view". If "monotechnic" view holds, cost-saving substitutions will be hard to realize without either reducing output or quality.

Let us, now look at the application of substitution in a health care production.
Substitution mean flexibility of replacing some inputs with others in a production process and still be able to realize the same output. But this does not mean that the two inputs are equivalent in any way but that alternative combination is possible (see Figures 3.9).

73

The figures show cases with no input substitution and one with input substitution. We have two inputs here, physician hours and nurse hours.

In the first figure, Isoquant q=1 show the fixed number of nurse hours and physician hours used in treating one patient. This is a case of monotechnic where production combine two inputs only. Adding either nurse hours or physician hours will not decrease hours of the alternative inputs. For example, moving N to the light (which mean increasing nurse hours) will not decrease the number of hours needed for the physician. The same applies to physician hours by moving up the vertical axis will not reduce the nurse hours. This is a case of fixed substitutability and it applies to where patient care requires certain professional task that only physician is trained and competent to perform. To use nurses to perform any physician task would result to less care or at least lower-quality care.

Figure 3.9 show how isoquant would look like if substitution was possible. The figure depicts a smooth downward-sloping convex isoquant curve. This demonstrates that different combination of inputs could be choosen without being wasteful unlike the fixed proportion case. Each point on the isoquant represents a different technique. The slope shows the rate at which nurses and physician hours can be substituted or exchanged despite maintaining the same output. The convexity of the isoquant implies that nurse and physician hours are not perfect substitute of each other. It implies there is an extent to which one can substitute these two inputs beyond which output it will not be possible. This is shown by the flat part of the isoquant curve at both tails. The relative flat part of the isoquant curve suggests that even if you add for instance nurse hours the physician hours will not change meaning neither will be the output. The converse is true for the physician hour's relatively flat isoquant part where nurse numbers remain more less the same with additional physician hours at the tail.

The downward sloping isoquant curve depict what is called diminishing marginal rate of technical substitution which shows that factor inputs are not perfect substitute of each other. The relatively flat part of the isoquant on both side of the curve is useful in

74

retaining the idea that while substitution is possible, it may be difficult, expensive or unsafe to have nurses do certain physician tasks , and vice versa.

Figure 3.9: Degree of substitution between physicians and nurses
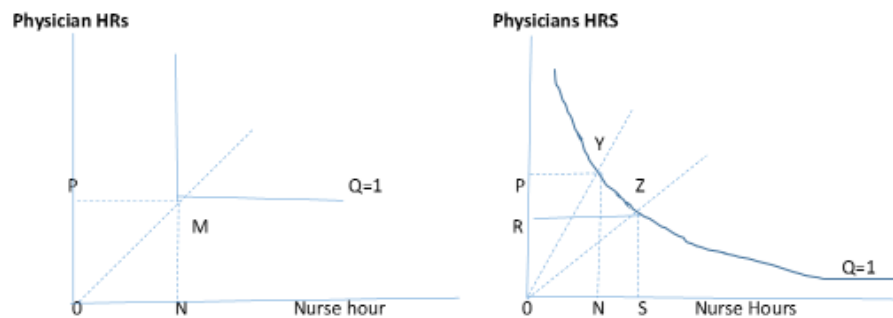
Elasticity of substitution in health care production:

Elasticity of substitution measures the possibility for substituting inputs which then measures the responsiveness of a cost-minimizing firm to changes in relative inputs prices. It is defined as

E_s= *percentage in factor input ratio/percentage change in factor price ratio*

This elasticity measure means that if a firm was a cost minimizer, then it would be responsive to changes in input prices and it would tend to respond by shifting away from

now costlier inputs to the now relatively cheaper inputs. The minimum value of *Es* is zero because it implies that the firm will the same input combination to produce a given level of output regardless of relative factor prices. Higher values of *Es* indicate a higher potential for substitutability.

Average cost

This is arrived in our example by dividing the total cost with the corresponding number of physician's visits. The firm experiences economies of scale if the average long-run cost is decreasing as output increases. The firm has diseconomies of scale if LRAC is increasing as output increases.

Given this scenario what output s would interest a profit-maximizing firm?—obviously the temptation would be to say Q_b where the LRAC cost is minimum. But this is not necessarily the case. A firm is not actually in the business of minimizing the average cost, it would do so if coincidentally the output that minimized costs also maximized profits.

Figure 3.10 Long-run Average Cost function
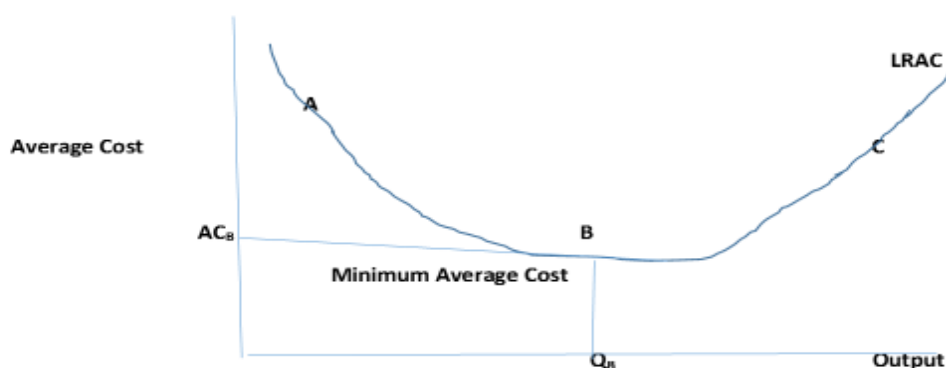
A related concept to economies of scale is economies of scope. Economies of scope are possible only for a multi-product firms and because many health care firms are

multi-product in nature this concept is highly relevant. Economies of scope occurs when it is possible to produce two or more different goods jointly and more cheaply than they can be produced separately.

Economies of scope: Think of a scenario where we have two hospitals one offering pediatric care (for children) and the other offering geriatric (care for the elderly). The question is would the total cost of pediatric and geriatric care be lower if one single hospital was providing both? It would be cheaper to combine the hospital and get scale economies but that is not our point for now. It might be cheaper to combine them if some inputs needed for both types of care interact well together. Possibly things learned in pediatric have some application in geriatric and maybe the two could support each other so that the result would be lower total cost. If this is the case, by producing the two different outputs jointly, we achieve economies of scope.

Lets use a cost function to show this;

We have two different outputs q1 and q2; q1 may represent pediatric care while q2 represent geriatric care hypothetically;

We want to produce output q1=100 and output q2=150, we can produce them either jointly or separately

Economies of scope exist if TC (q1=100, q2=150) is less than TC (q1=100,q2=0) +TC(q1=0,q2=150)

Hence economies of scope exist if the following inequality holds

TC (q1=100, q2=150) < TC (q1=100, q2=0) +TC(q1=0,q2=150)

Hospital cost Analysis

Hospital cost analysis require consideration of their heterogeneous hospital output. They differ by type of case they treat what is called the case-mix problem. Hence a multi-product approach in cost analysis is deemed appropriate by many health economists. Though this method is deemed superior, multi-product approach rarely incorporate more than four or five hospital output categories. We can illustrate a case – mix issue by use of Figure 3.11.

The figure illustrate a case where the long-run average cost function curve is flat exhibiting neither economies of scale nor diseconomies of scale. The cost curves represent three hospitals, dealing with uncomplicated cases, more complicated cases and most complicated cases respectively. The three horizontal lines represent long-run average cost curve for each hospital type. The difficulty with analysis is that the estimated LTAC for all the hospital might appears like it is rising thus mistaking it with diseconomies of scale. But looking at the case-mix each hospital has got its own LARC and comparing these hospitals in terms scale economies will require very a lot of data most of some of which are not observable. The scale economies could be affected by quality issues which means that in some hospital quality could dedicate the price while in others capital investment could have an effect on the long- run average cost.

Figure 3.11: Hospital cost Analysis
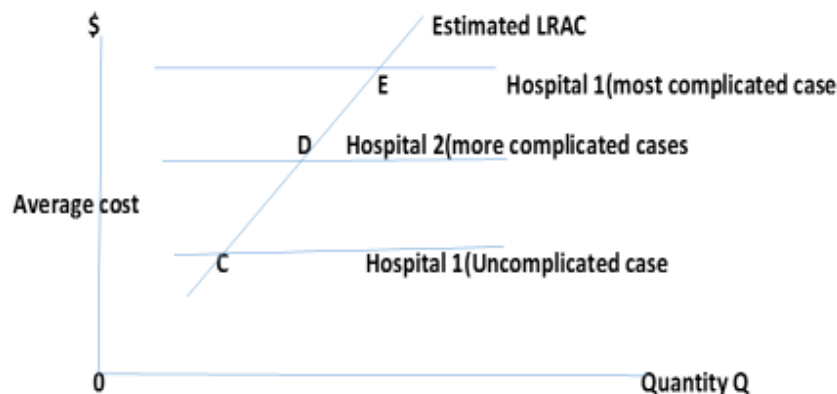
Empirical findings on the extent of economies of scale in hospital services

The theoretical relationship between hospital cost and size is the U-shaped long run average cost curve. Other things being equal, the average hospital cost may be expected to decline initially and then rise as size is increased. How fast the gains and losses from scale of operation occur is what determines the shape of the long- run average cost curve.

Figure 3.12 Hospital Size and average cost

The relationship between hospital size and average cost
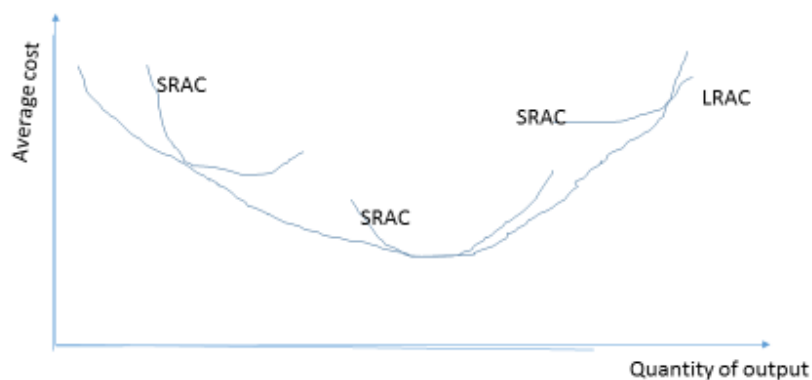
This theoretical relationship is affected by the fact hospitals are not homogenous in size or other characteristics, they are multiproduct firms. They produce both inpatients and outpatient services which differs in quality, type of patient treated, severity of illness etc. Attempt to measure the effect of size on average cost has been affected by the inability to hold all the factors that affect hospital costs while estimating the net effect of size alone. The result of failing to hold constant differences in costs between hospitals as a result of seriousness in illness is demonstrated in the following

Figure 3.13 Average cost between hospital

Variations in Average cost between hospitals
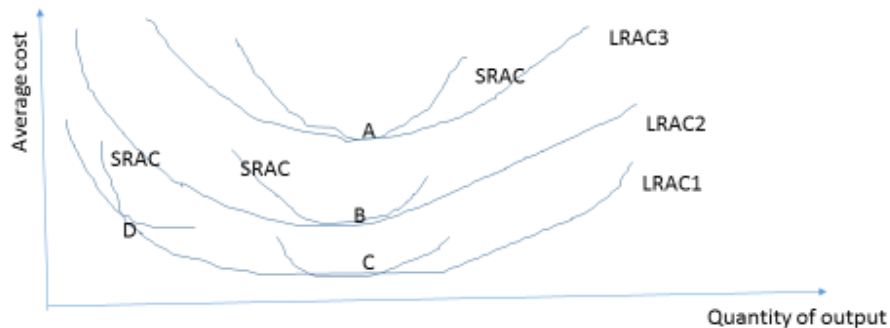
The figure show that hospital C could be providing fewer services per patients than A and B yet register the same number of visits. It on a lower long-run average curve. The difference is due to the type of product/service. It also appears like hospital D which is smaller has lower cost than hospital A which is of a large size. Notable is the fact that hospital A is producing the product at the least costly facility for it is producing at minimum point of its LRAC3. Hospital D on the hand is producing a product that is least costly and at a facility that is not least costly.

Why is this so?

The perception is that large hospitals are more costly than smaller hospital. This could be due to the differences in how hospitals pay physicians (large hospitals has more salaried physicians, interns and residents than smaller hospitals. An important limitation of most previous hospital cost studies is the failure to adjust hospital cost for the differences among hospitals output. The empirical approach used to derive the effect of size on cost, while controlling for other differences is to use a multivariate analysis:
That is,

AC=f(B,S,C, Q, V, P, D, O);

AC = average cost

B=hospital size

S=hospital service capability measured in terms of enumeration of facilities and services in hospital

C=patient mix, proportion of patients in a given number of disease classification

Q= measure of quality, e.g. inputs per patients

V=severity of illness, e.g. number of surgical procedure

P=adjustment for differences between hospital for wages and other factors prices that are unrelated to hospital size

E=difference in hospital efficiency that are unrelated to hospital size

D=education programs (e.g. number of interns, affiliations with a medical school, nursing home etc.)

O= other variables such as the physicians contribution and other hospital output such as outpatients visits

Medical technology and healthcare productivity

There has been a rapid pace in technology in healthcare industry. This advancement raises economic questions about the effects this changes will have. What is known is that technology can either lower costs when it improves productivity of healthcare resources or it might increase costs when it improve the quality of care or introduces new and costlier products.

The assumption that it increases cost in the health sector is the reason that many health economist hypothesis that it is a key driver of health sector inflation. Widespread and deep insurance coverage has also been hypothesized as a reason why technological innovation in health sector has been on rise. Key aspect to not is that the effect on cost and quality improvement cannot happen unless technological adoption or diffusion has penetrated to the providers.

Answer!

Technological change necessarily entails an improvement either by providing less costly production methods for standard old products, or alternatively by providing new or improved products. Generally, it will be less costly to produce a given output holding quality constant according to cost experience. However, the mix of products and services sold may change the direction that raise the average cost of a patient day, a case treated, or a physician visit. Hence the total health care expenditure per capita may rise.

The concept is illustrated in our figure below. Figure A show a state of a decreasing cost after the introduction of medical technology. In figure A the quality of the care is unaffected. Before the diffusion of technology, the isocost for treating 100 cases was at E. After the introduction of the technology the isocost moves to E'. It costed much higher to treat 100 cases before the adoption of medical technology.

Figure B on the other hand show a scenario where the introduction of technology has caused treatment of 100 cases to be much higher than before the adoption of the medical technology. This is shown by comparing isocost at F (before technology) and F' (after adoption of technology). The implication is that under figure B, there is an improvement in quality which the provider charges some extra cost to treat the same medical condition or to treat the same number of patient.

Figure 3.14 Technology Adoption in Medical technology
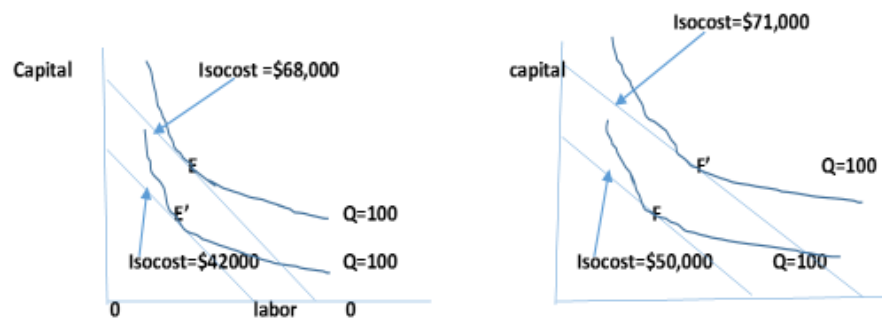
Diffusion of new health care technologies.

It takes time before a new innovation is widely adopted by providers. Some firms adopt rapidly some slowly and some not at all. There two main principles that guide provider or adopters. One is the profit motive while the other one is information channel. Physician will tend to adopt a new for example surgical technique if it is expected to increase their revenue base –which could arise through enhancing their prestige or by improving the well-being of their patients.

The information aspect is derived from sociology in what is called compatible principle which emphasis the role of friend colleagues, journals and conferences at informing and encouraging the adoption decision. Escarce (1996) emphasized the information externalities inherent in adoption by the first physician to adopt. An externality is the uncompensated, beneficial effect on the third person caused by action of the market who in this case is the adopter. After adopting the technology, the physician communicates to friends and colleagues the expectation that the new product will benefit his or her patients and practice. The prices tend to build on itself until the all the potential adopters has acted and community maximum potential is attained.

The adoption occurs slowly at first, then at an increasing rate that continue at decreasing rate asymptotically approaching its limit as shown in Figure 3.15 which illustrates the classic pattern of diffusion of diffusion as a logistic curve.

Readings: Escarce 1996: Externalities in Hospital and physicians adoption of New surgical Technology: An Exploratory Analysis Journal of health economics 15(1996) 715-734.

Figure 3.15 The diffusion of new technology
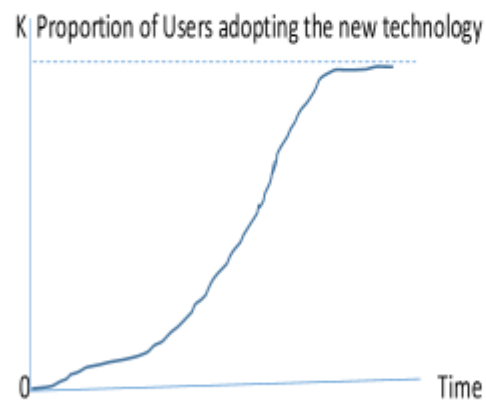
Efficiency in Health care Production

Concept of economic efficiency is related to both the demand and supply side of the industry as much as efficiency studies are concerned with optimal output. Economic efficiency in health care demand is related to economic efficiency in supply of medical care through prices. In that case, policy recommendation geared towards improving efficiency of the market should center on factors causing rise in the cost of medical services. A decrease in the medical cost is also seen as a "return" or saving especially in an environment where expenditure are rising over time. They affect optimal utilization

of medical services as well as the consumer surplus. The policy implication of economic efficiency of the supply side of the medical care can be assessed through supply side**. HOW?** If the supply side of the medical care is relatively inelastic, it will require relatively large price increase to bring forth an increase in medical care outputs. This is because the providers are not attempting to minimize their cost

- This will affect the re-distribution programs proposed on the demand side especially the type of national programs that can be instituted.
- This can be demonstrated by the following figure

Figure 3.16 Efficiency in health services
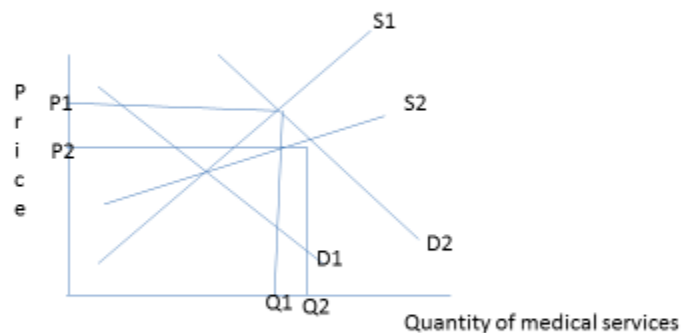
A relatively inelastic supply curve, represented by S1 would, with an increase in demand from D1 to D2 result in a greater price rise and a smaller increase in service provided than if the supply of medical services were more elastic. We need to note that a more elastic supply schedule would , for the same increase in demand give Q2-Q1 should be more services at smaller increase in price i.e.,P2 rather than P1.Total cost of

the increase in demand would be P1*Q1 in the elastic case versus P2*Q2 in the situation where supply is more elastic.

In the case of relatively elastic supply curve, more of the expenditure would go for increased medical services unlike in the relatively inelastic supply where more will go to cater for rapid increase in price with smaller increases in medical services. Key implication is that the cost of a national health program would be greater and the availability of services diminished when supply is more inelastic. This is due to the fact that, the greater cost of instituting such a program as well as it comprehensiveness is reduced. This then implies that the economic efficiency of the supply of medical care affect the type of national health insurance program that is developed. It also influences when these programs should be implemented and on what services or coverage.

It also has a redistributive effect among different population groups in a society in that greater inelasticity will mean greater increase in prices, wages and income of the provider of medical services. Notably, the financing will be done by the population in form of out-of-pocket, insurances premiums and taxes. Hence there is need to assess the causes of increased prices in the supply of the medical services. If the analyses reveals that supply of medical series is solely due to the nature of production function and that providers are attempting to minimize their cost of production, then very few changes will be possible.

Hence policy wise more demand will imply higher expenditure at less additional medical services as seen earlier. If on the other hand production function is artificially constrained by say legal, it implies that some providers have market power that would make them increase prices. This legal restriction will also provide no incentives for the said providers to minimize the cost of production. In this kind of a scenario it should be possible to improve the performance of the medical sector.

Types of Efficiency in health sector analysis

Efficiency in health sector is assessed through three type of efficiency measure viz

1. **Technical efficiency:** Exists when production occurs on the technical frontier—operationally, when no greater output at a given level of quality can be achieved with a given set of resource inputs.

2. **Productive efficiency.** Exists when the various factors of production are used in optimal combination—operationally, when no greater output of specified quality can be achieved at a given resource cost.

3. **Allocative efficiency:** Exists when resources are allocated so as to maximize the benefit to a population—operationally, when no alternative allocation of resources can make someone better off, without at least one person being worse off.

These terms are theoretically meaningful in health care but applying them is hazardous given the nature of the health care "product." Robert Reischauer pointed out that it is possible to evaluate the efficiency with which steel is produced, because inputs (labor and capital) can be measured carefully, (true) input prices can be estimated with confidence, and quality (tensile strength, purity) can be specified with a high degree of precision.

But "health care is not steel." In health care, the factors of production are rarely captured accurately, and the prices of those factors are highly distorted by insurance and cross-subsidization of health care service lines. Even more important, though, is the terrible complexity of standardizing the "health care product"—the nature and outcome of care so often depend upon the characteristics of the patient that evaluation of technical and productive efficiency is difficult at best.

Let' discuss each briefly
TECHNICAL EFFICIENCY
Technical efficiency occurs when the firm produces the maximum possible output from a given set of inputs. This efficiency measure must be "pareto" efficient. Pareto efficiency means a situation in which it is not possible to make one better off without

87

making another worse off. This type of efficiency suggests that if an intervention can produce same or greater outcome with loss of one type of input while holding the other input constant – then this intervention is technically inefficient. For instance, if you can treat 10 people using 2 doctors and at the same time treat 10 people using one doctor then it means that this type of health care program is technically inefficient.

Situation in which health care interventions for a particular health states are each performed with least amount of inputs. Technically efficient should only have one outcome if maximization principle is attained. Technical efficiency specifically refers to the optimal combination of inputs meaning using minimum combination of labour and capital to produce a certain quality of goods or outcomes. Technical inefficiency on the other hand implies that the producer is not achieving maximum output from a given input combination.

The danger with emphasis on technical efficiency is that it can lead to suboptimal recommendations. This is because technical efficiency deals with intermediate outputs as a measure of effectiveness. Effectiveness measure in health care is measured through health outcomes and not through intermediate output or through input. For example, technical efficiency will measure – waiting time, hospital visitation, number treated, bed occupancy etc. If number treated is used as a measure of efficiency, it might not give a clear picture of the health outcome in terms of how many recovered or how many life are saved out of this intervention. Thus, it is one thing to measure efficiency and another thing altogether to measure effectiveness. It is in this length that economic evaluations should focus on final health outcomes as much as focusing on technical efficiency.

Demonstration of Technical Efficiency
F(L) is production possibility frontier in figure 3.17. PPF show the maximum of quantity that can be produced by each amount of inputs this production possibility curve is generated in health care setting through a method data called Date Envelopment

Analysis (DEA). DEA method establish maximum output that is produced by some combination of inputs factors. From our Quality frontiers firms 1,2,3, are inefficient because they are of the frontiers. They are using more inputs to produce low level of output using same labour that other firms are using to produce more output.

A firm's inefficiency is measured as relative distances from the frontiers. The distance focus could be either horizontal or vertical distance. Horizontal is asking how many fewer inputs could have achieved this output. The vertical distance is asking what output could have been achieved with these inputs.

Figure 3.17 Technical efficiency
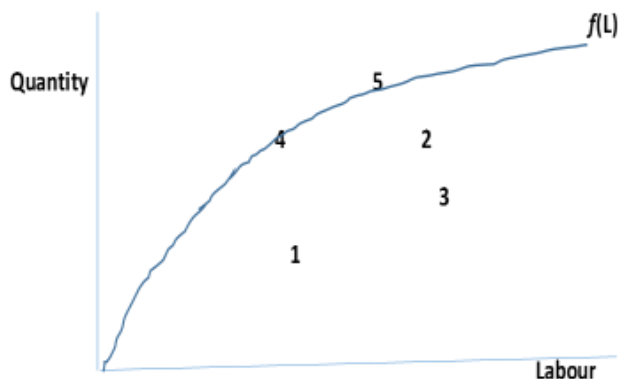
Productive Efficiency

Productive efficiency means that, given the available inputs and technology, it's impossible to produce more of one good without decreasing the quantity of another good that's produced. All choices along the PPF in Figure 3.18, such as points A, B, C, D, and F, display productivity efficiency. As a firm moves from any one of these choices to any other, either health care increases or education decreases or vice versa. This makes sense if you remember the definition of the PPF as showing the maximum amounts of goods a society can produce, given the resources it has. Thus, producing efficiently leads to maximum production, which is what the PPF shows. However, any

89

choice inside the production possibilities frontier is productively inefficient and wasteful because it's possible to produce more of one good, the other good, or some combination of both goods. Wasting scarce resources means the society is not producing as well or as much as it could, so it is not operating on the PPF.

Using the PPF principle, it is impossible to produce more of one good without producing less of another. It is also defined as situation in which the economy or an economic system (e.g., a hospital, an industry, a country, etc.) could not produce any more of one good without sacrificing production of another good and without improving the production technology. The same principle applies when deciding on the input mix. According to productive efficiency it is not practical to use more of one factor input without reducing the other factor input. If one has to stay at the same level of quantity. For example, point R is productively inefficient because it is possible at choice C to have more of both goods: education on the horizontal axis is higher at point C than point R (E_2 is greater than E_1), and health care on the vertical axis is also higher at point C than point R (H_2 is greater than H1).

In health care – The Concepts of productive efficiency refers to the maximization of health outcome for a given loss of the minimization of cost for a given outcome. Unlike Technical efficiency which cannot directly compare alternative interventions, where one intervention produces the same (or better) health outcomes with less (or more) of one resources and more of another, productive efficiency enables assessment of relative value of money of interventions with directly comparable outcomes. The only thing that it cannot address is the impact of reallocating resources at a broader level – e.g., for cancer care to mental illness – This is because the health care outcomes are not comparable.

Figure 3.18 Productive efficiency
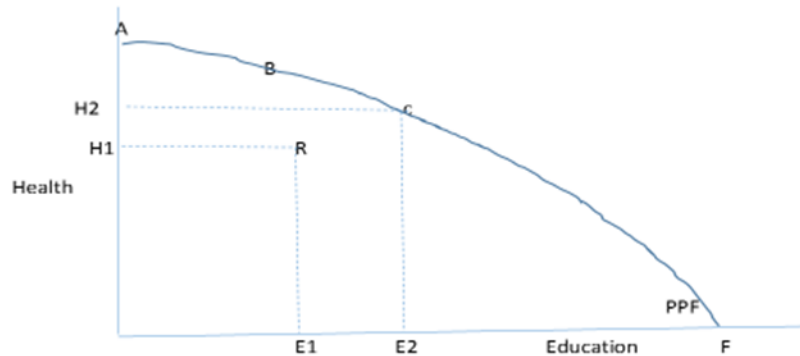
ALLOCATIVE EFFICIENCY

The concept of a locative efficiency takes account not only of the productive efficiency with which healthcare resources are used to produce health outcomes but also the efficiency with which these outcomes are distributed among the community. This allocative efficiency spells out situation in which health service resources are put to their best possible use so that no further improvements in the health status of the community as a whole are possible. In a sense a locative efficiency is achieved when resources are allocated to maximize the welfare of the community. It is more concerned with distribution and allocation of resources in society. Allocative efficiency has a key attribute of distributive efficiency which occurs when good and services are received by those who have the greatest need for them.

In economic sense allocative efficiency is the market conclusion whereby resources are allocated in a way that maximizes the net benefit attained through their use. Thus, it refers to a situation in a county where limited resources of a country are allocated in accordance with the wishes of consumers.

91

The concept of a locative efficiently in health care takes account not only of the productive efficiently with which healthcare resources are used to produce health outcomes but also the efficiently with which these outcomes are distributed among the community. In technical language allocative efficiently addresses the issue of input prices as much as it takes care of technical efficiency.

Figure 3.19 Allocative efficiency
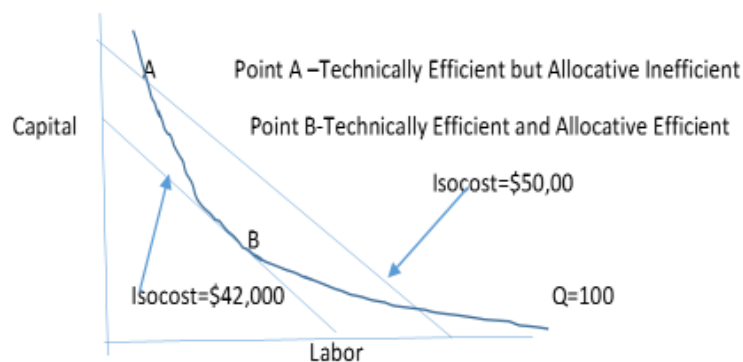
In the figure 3.19, the isoquant for 100 cases treated is illustrated as the curve labeled Q=100. At point A 100 cases are treated but at higher cost of 50,000. At point B same number is treated but at a lower cost of 42,000. Thus, point A is Productive and technically efficient but allocative inefficient.

Point B is both technically efficient and allocative efficient.
The reasoning is that the Resources saved under condition B could be used to benefit community in many other ways that could relate to improving health outcomes.

One can also look at Allocative efficiency in terms of **normative perspective**. This where allocative efficiency implies that the particular mix of goods a society produces

92

represents the combination that society most desires. For instance, a society with a younger population can be deemed to have a preference for production of education, over production of health care. Under this circumstance If the society is producing the adequate level of education that the society demands, then the society is achieving allocative efficiency. This is a normative issue that could be a controversial question.

Thus, in a health care system allocative efficiency could mean government supplying the quantity of resource inputs to both preventive and curative care that residents demand. Only one of the productively efficient choices will be the allocative efficient choice for society as a whole as much as there will be many points in the production possibility frontier. For example, form our example on health and education, in order to achieve allocative efficiency, a society with a young population will invest more in education. As the population ages, the society will shift resources toward health care because the older population requires more health care than education. Thus, allocative efficiency can take a dynamic angle.

Empirical technique to estimating efficiency in healthcare.
Several different empirical techniques have been appeared to address efficiency assessment. These techniques can be grouped into non-frontier and frontier empirical studies. In the non-frontiers assessments, actual output or firm costs for two or more groups of firms are compared while attempting to control for the effect of extraneous variables, (not independent **variable**, but could affect the results of the experiment). For the frontier studies actual output or firm costs are compared to the best possible experience.

Frontiers assessment are inclined to the technical and allocative efficiency. Two type of empirical analysis that have emerged includes Data Envelopment Analysis (DEA) which is deterministic in nature and Stochastic Frontier method (SFA) which provides an alternative that allows for incorporation of both inefficiency and measurement error. DEA finds the frontier isoquant for selected level of output (for example, Q=100) by

forming an envelope of data. Researchers use linear programming to construct this efficiency outer shell of the data points. See figure 3.20 . DEA is said to be nonparametric in terms of it analysis because it imposes no assumptions about parameters of the underlying distribution of inefficiency.

On the other hand SFA treats each firm uniquely by assuming it to be affected by a potential shocks to its ability to produce care. Hence unlike DEA, SFA must guess in advance the statistical distribution of the inefficiencies.

More work on DEA/ SFA to be done as presentation in class through group work

Figure 3.20: DATA envelopment method
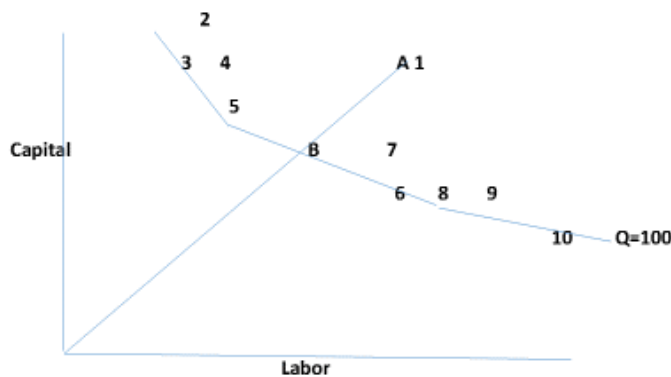
ASSIGNMENT

a) Discuss major causes of inefficiencies in a health care set up

b) Discuss frontier method used in measuring efficiency and show their comparative strength and weaknesses

c) "In an environment in which one sector's cost is another sector's revenue, there seems no way to reduce total costs without someone's income falling". Discuss

94

the relevance of this statement under efficiency in health care or what you might want call a zero sum game.

Key References

Folland, S., Goodman, A. and Stano, M. (2016). The *Economics of Health and Health Care*. Pearson International Edition Seventh Edition. Routledge. ISBN-10: 1292020512 or ISBN-13: 978-1292020518

Rice, T.H. and Unruh, L. (2015). *The Economics of Health Reconsidered*. Chicago: Health Administration Press 4th Edition

Feldstein, P.J. (2012) *Health Care Economics*, 7th Edition, Cengage Learning Publishers Inc. ISBN-10: 1111313261 | ISBN-13: 9781111313265

Witter, S., Ensor, T., Jowett, M. and Thompson, R. (20015. *Health Economics for Developing Countries: A Practical Guide*. KIT Publishers, 2015. ISBN-13: 978-9460221316 ISBN-10: 9460221319

Chad D. Meyerhoefer , David E. Sahn & Stephen D. Younger (2007)
The joint demand for health care, leisure, and commodities: Implications for health care finance and access in Vietnam. *The Journal of Development Studies* Volume 43, Issue 8

Gertler, P., Locay, L. and Sanderson, W. (1987) Are user fees regressive? The welfare implications of health care financing proposals in Peru. Journal of Econometrics, 36(1): 67–88

Sahn, D. E., Younger, S. D. and Genicot, G. (2003). The demand for health care services in rural Tanzania. Oxford Bulletin of Economics and Statistics, 65(2): 241–260

Deininger, K. and Mpuga, P. 2005. Economic and welfare impact of the abolition of health user fees: evidence from Uganda. Journal of African Economies, 14: 55–91. [Crossref], [Web of Science ®]

Carrin G. Strategies for Health Care Finance in Developing Countries, with a Focus on Community Financing in Sub-Saharan Africa New York St. Martin's Press 1992 [Crossref] , [Google Scholar]

Dor, A., Gertler, P. and Van Der Gaag, J. 1987. Non-price rationing and the choice of medical care providers in rural Côte d'Ivoire. Journal of Health Economics, 6(4): 291–30

Dow, W. 1995. "Unconditional demand for curative health inputs: does selection on health status matter in the long run? RAND". Labour and Population Programme Working Paper, No. 95–22 [Google Scholar]

World Bank. 2001. Vietnam Development Report 2002: Implementing Reforms for Faster Growth and Poverty Reduction, 4–5. Washington, DC: World Bank. [Crossref] , [Google Scholar]

Cutler, D., Gruber, J., 1996. Does public health insurance crowd out private insurance? Quarterly Journal of Economics 111, 391-430.

Bahar Bayraktar Sağlam (2014) Private Health Care Versus Public Health Care: Complements or Substitutes. Department of Economics, Hacettepe University, Ankara, Turkey e-mail:sbahar@hacettepe.edu.tr

Jeanne S. Ringel Susan D. Hosek Ben A. Vollaard Sergej Mahnovski The Elasticity of Demand for Health Care A Review of the Literature and Its Application to the Military Health System. Prepared for the Office of the Secretary of Defense National Defense Research Institute RH e a l t h RAND Approved for public release; distribution unlimited

John Wildman and Peter McMeekin. (2014).Health care and social care: complements, substitutes and attributes. Economics, Newcastle University, IHS, Newcastle University

4 Markets and Market Failure in Health and Health Care
Learning objectives

By the end of this chapter the learners should be able to:

a) Demonstrate the application of market structure frameworks in the market for health care

b) Examine the role of market imperfection in the determination of prices in the health care market.

c) Evaluate how comparative static affect price mechanism in the health care mark

d) Examine the characteristic of healthcare and services

e) Describe a public good and provide example from a healthcare good perspective

f) Demonstrate with the reliance of healthcare good in the domain of a public good

g) Discuss different types of externalities and provide their application in healthcare markets.

h) Evaluate the causes of market failures in healthcare market

i) Assess policy mechanism that can be put in place to solve market failure.

j) Argue the role of government in healthcare

k) Synthesis the operation of pharmaceutical industries and health workforce in terms price settings and regulations the market for healthcare

Health Care Markets

A market is a system in which in which the mechanism of variable prices functions freely to signal consumer preferences and to encourage resource allocation so as to satisfy those preferences. **Healthcare market** is basically the terms related to the medical field, the product or services which are used by a medical stream person or for someone's **health** contributes to **the healthcare market** where it can be products or services of pharmaceutical or personal **care** or hospital management resources or biotechnology.

Why Study Healthcare Market?

Society values greater efficiency, technological progress and fairness in the distribution of income. Healthcare market is key area of concern to health economist in the sense that these three economic issues need attention as well in this market:

- If markets do not produce desirable outcomes, public policies can be aimed at correcting market failures
- Competitive market: Profit-seeking firms are usually driven by competitive forces to serve the interests of society by efficiently allocating scarce resources
- When market power is present, profit-seeking may lead to a misallocation of resources

Table 4.1 summarizes features of various market structure that guides at categorization of our healthcare market. Generally, we have two types of market structures that, is, Perfect competition and imperfect competition. Imperfect competition is sub-divided further into, monopoly, monopsony, Duopoly, oligopoly, and monopolistic

Table 4.1 Market structure

Market structure

Characteristics	Perfect competition	Monopolistic competition	oligopoly	Pure monopoly
Number of sellers	many	many	Few, dominant	one
Individual firms market share	tiny	small	large	100%
Type of product	homogenous	differentiated	Homogenous or differentiated	Homogenous by definition
Barriers to entry	none	none	substantial	complete
Consumers information	perfect	Slightly imperfect	Perfect or imperfect	Perfect or imperfect

Perfect competition

Under the "general equilibrium theory" of economics, a perfect market is defined by several conditions, collectively called perfect competition. While the following is not a complete list of

Conditions, here are the characteristics of a perfect market that are most pertinent to the analysis of health care markets:

1. There are large numbers of buyers and sellers of a homogeneous product. Here the assumption is that none of them can affect price by their behavior. All they are price takers. Some healthcare markets are more competitive than others, in terms of number of players. For example, in rural areas, they may be only supplier of a particular service in that locality.

2. All consumers and producers have complete knowledge of price, quality and other aspects of a good or service (e.g., in health care, aspects such as need and efficacy).This mean that the suppliers have access to the knowledge about the good and the consumer know the price being charged for the good by each supplier and able to choose the lowest price. As noted earlier this is not possible in

the health care due to information asymmetry between the physician and the patient

3. Buyers are rational in that they only purchase goods or services that increase their "economic Utility" (or "satisfaction"). This condition is not likely to hold in the health care market because a sick person in most cases is not able to make rational decision for in some instances, he or she will rely on somebody else to take them to hospital.

4. There are no barriers to entry or there is simply free entry. This means that if there is excess profits to be made, then new firms can start to produce and sell that good, bringing profit down to minimum level. Consumers are then paying no more than is absolutely necessary to cover costs of production and have more resources available to satisfy other needs. Free entry is rarely found because of regulations, registration and licensing intended to protect consumers from the 'quacks'-people who offer cures but lack essential training and skills.

5. In large part because there are large number of buyers of a homogeneous product who are rational consumers with complete knowledge, providers do not have the power to set prices or influence demand. In healthcare this is not likely to hold because for one the treatment is not homogenous, patient have no knowledge of the treatment option and as mentioned earlier they are not sure of what they of the price they are paying in order to make a rational decision on whether to buy or not.

6. There should be constant (or decreasing) returns to scale in production. This means that bigger producers should not face lower costs per unit than small producers. If they do, then they will be a tendency towards monopolies as large producers undercut and drive smaller producers out of the market. In secondary or hospital there is some evidence that economies of scale or scope exist implying that they are able to generate savings through pooled resources or more efficient use of staff.

When these and other conditions of a perfect market hold, it can be shown (graphically and mathematically) that a market will reach an equilibrium in which the quantity supplied for every product or service equals the quantity demanded at the current price.

This equilibrium is called a "Pareto optimum," meaning that nobody can be made better off by exchange without making someone else worse off. Through further theoretical analysis, it can be further shown that if the assumptions of perfect competition hold, the overall welfare of a society is optimized

But is this the economic reality when it comes to health care market?

Intuitively it is clear that many of the above assumptions of perfect competition do not hold in the field of health care due to various complicating issues. In the healthcare market we have the supply side that comprises the firms that produce commodity called healthcare (doctors, nurses, hospitals nursing homes professionals etc.) while on the demand side we have consumers which in medical terminology are the patients. Supplier would be thought as pricing their goods in what is called reimbursement and consumer side paying through out-of-pocket expenditure for instance.

In this context we have sub-markets that the break the assumption of a perfect competitive market. These sub-markets include:

- Institutional care markets
- Market for health manpower and related factors
- Health professional market

These three markets comprise our health care markets. It is there interaction which determine the quantity produced as well as the demand of health care. Each of this market has got its demand and supply side. In this market the patient is the initiating factor after a decision to get in touch with the medical market. Assuming that he or she make a decision to avail himself or herself to general practitioner (GP), the GP will assess him or her and decide on how much treatment to administer. This decision will be based on both economic and non-economic consideration.

Given the continuum of market structure in the Table 4.1, healthcare market fail to fall the perfect competitive market classification due it's the limitations linked to most of it characteristic. The **market** for health-care services is considered an **imperfect market**

because -- 1) Health care is a heterogeneous product, as the patient can experience a range of outcomes; 2) Patients who are insured have third-party payers covering their direct medical expenses; and 3) A "**market** price" is lacking that is no feedback mechanism exists that reflects the value of the resources used in health care 4) only supplier has complete knowledge of price, quality and other aspects of a good or service. Yet the "normal" economic model presumes that consumers are fully informed about prices and quality of every good available in the market, and that all producers have complete knowledge of available production technologies.

Key implication of the model, for example, is that under the assumptions of perfect competition the price paid by consumers is determined by the intersection of independent supply and demand curves. In health care, however, the price paid by the consumers is *not* the full price especially with insurance; rather, when patients check out after a doctor visit or hospital stay, the price they typically pay is the "co-pay" as defined by their insurance plan. And since this co-pay (i.e., the actual price paid by patients) is much less than the full price paid to the provider by the insurance plan, it follows that the utilization of health services by patients will be much greater than if they were facing the full price.

Another problem with applying the market competition model to health care is that demand is not necessarily independent of supply. The assumption of perfect information, combined with rational decision-making and a large number of buyers and sellers, leads to independent supply and demand curves. Put differently, when consumers know all there is to know about a service, those who supply that service cannot influence demand. But patients do *not* know all there is to know about a particular medical condition, diagnostic test, imaging study or treatment. Therefore, the providers of health care, who make recommendations regarding visits, diagnostic tests, imaging studies and medical or surgical treatments, can indeed influence a patient's decision to purchase such services.

Further, a significant market imperfection in health care services is that providers of healthcare *can* influence demand due to imperfect consumer information. And because of widespread Insurance coverage — whether government or commercial — that significantly reduces the price faced by consumers, the constraint on utilization that would normally be achieved by price does not tend not apply in healthcare market.

Comparative statics - shifts in demand and supply curves

Comparative static analysis study the change in the quantity demanded of a particular goods as a result of changed conditions. For example, what would be the effect of change in income, change in price of the good, change in price of substitute good on the demand of a good or a service? One want to compare the difference between the two levels of equilibrium. Of our relevance here is how does the market react to events that influence the demand for or supply of medical services? Recall that changes in factors other than output price will cause the demand or supply curve to shift. An increase in consumer income will cause the demand curve for physician visits to shift to the right. An increase in the wage of nurses will cause the supply curve for hospital stays to shift to the left. These shifts in the demand or supply curves will lead to a change in equilibrium price and quantity

Key to note: Predicting such changes is referred to as *comparative static analysis*

In order to clearly how comparative static work, we will use the effect of AIDs in the demand for gloves. We will proceed to see how output is affected with the assumption of competitive market model. In the mid-1980s, the AIDs epidemic led to an increase in the demand for latex gloves among health care workers. The epidemic led to a shift to the right in the demand curve for latex gloves.

As seen in figure 4.1 excess demand for gloves developed, leading to a temporary shortage of gloves. The shortage of gloves led buyers to bid the price of gloves upwards. As the price bid for gloves rose, sellers increased their quantity supplied of gloves. This process continued until a new short-run equilibrium was reached.

Figure 4.1 comparative Statics

Comparative Statics (Long run)
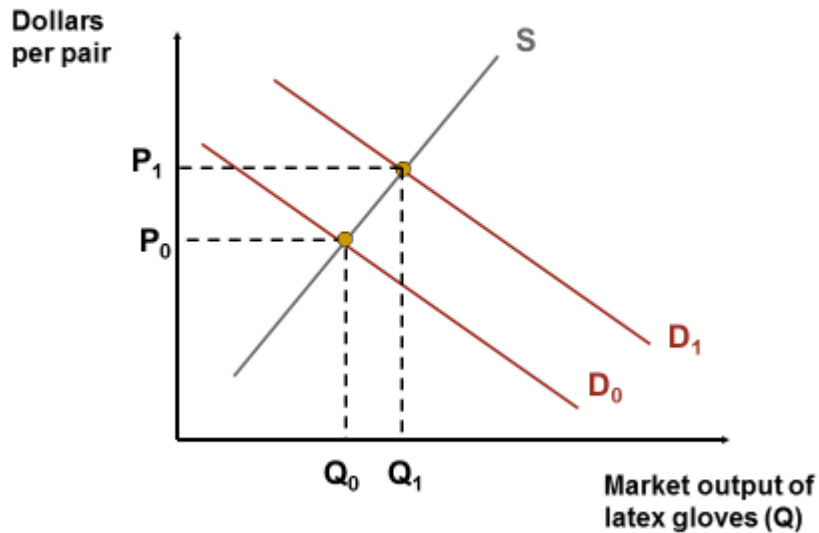
Before the epidemic, each glove maker was earning 0 profits because the firms are assumed to be operating under perfect competitive market structure. The increase in equilibrium price after the epidemic implies that all glove makers are earning positive profits as shown, MR>ATC.as seen in figure 4.2

$$\pi = (P_1 \times Q_1) - (Q_1 \times ATC(Q_1))$$

Figure 4.2 Positive profit of a perfect competitive firm

Comparative Statics (Long run)
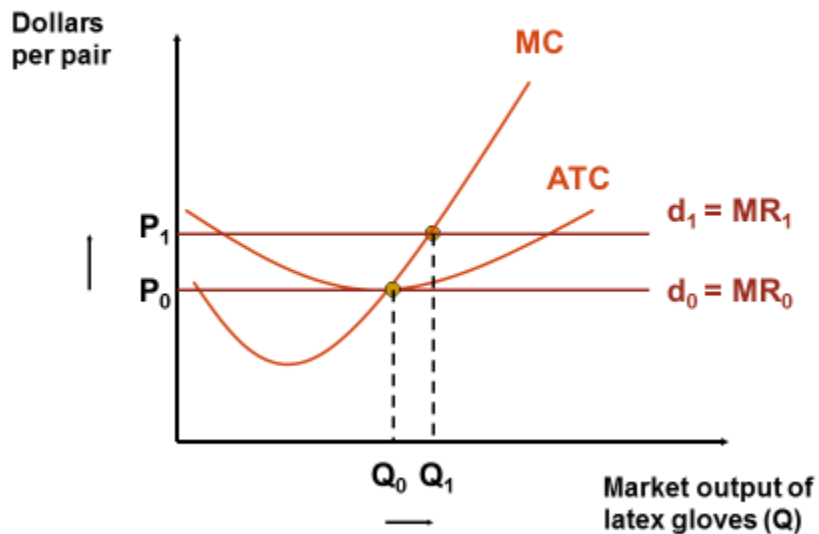
Due to these abnormal profits, other medical suppliers made plans to build new manufacturing plants to make gloves, in the hopes of making profits. Entry of the new plants into the market increased the supply of latex gloves in the long run causing the supply curve for gloves shifted out as shown Figure4.3

Figure 4.3 Comparative statics

Comparative Statics (Long run)
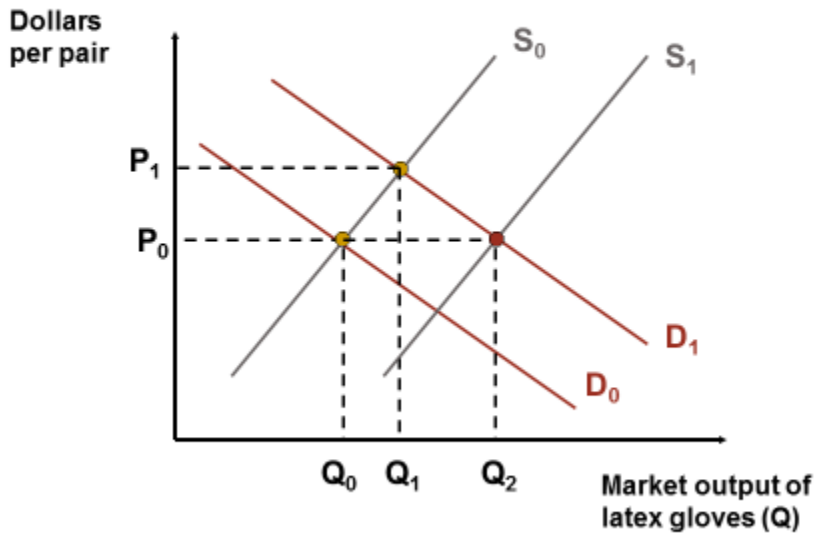
As the supply curve for gloves shifts out, the price of gloves begins to fall. Note that the quantity of gloves sold on the market also increases due decrease in price. As the price of gloves fall, profits also fall. The process continues, until the price of gloves falls back to P_0, where profits for all glove makers are again equal to 0. See Figure 4.4 arrow pointing down.

Figure 4.4 Positive profit 2

Comparative Statics (Long run)
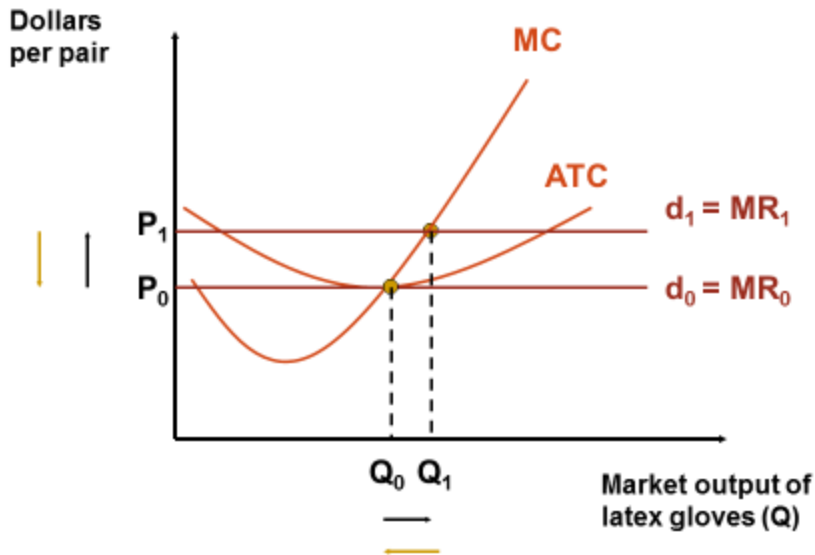
Assignment

a) A policy maker suggests increasing sick pay, to be financed by deduction from the wage of employed people. What would be the effects on the population's health behavior and health status.

b) One can derive demand functions for health and medical services from the Grossman model. Explain the difference between the two demand functions .why does demand for health and for medical services depend on age?

Characteristics of healthcare goods and services in the health care markets

Health care commodity is different from other goods and services in the in the traditional market. This is because of it features which are somehow unique in the demand and supply side. This section discusses the feature that have made health care good and services to depart from our usual daily known well and service.

a. **Presence and extent of uncertainty**; Consumers are uncertain on their health status and need for health care at any coming period. This implies that demand is irregular at any given time to both individual and health care provider Unlike economic analysis consumers do not know the expected outcomes of various

108

treatment likewise the provider cannot predict the outcome of treatment with certainty. This uncertain in health sector suggests that government has to intervene in order to rectify market failures that which is likely to affect poor whose saving is limited if not absent to deal with uncertain future sickness that might befall then. The role of government is then very critical in the provision of health care.

b. **Information asymmetry.** Information symmetry implies that perfect market where goods and services are bought and sold, both the buyer and the seller have correct information about the goods in the market. In a market where this is not the case meaning that one player has more information about the good or service then the market faces information asymmetry. In health care market, the consumer or the patient lacks information about type and amount of treatment required and its effectiveness. The main reason being that the consumer lack medical knowledge which is that require complexity of the technical data. Tied to this there is multiplicity of choices that the patient has no knowledge upon. The distressed state of mind of the sick person also contribute to the ability to make the right decision on what medical to demand from health the provider. Give that the suppliers knows the characteristic of good they supply they more likely to exploit the consumers in a free market environment. Consumer can end up getting more of the health goods than they require or even low quality health good.

c. **The agency relationship**. Agency relationship is a situation where one entrust somebody else to perform a duty on his or her behave. In health care, given the nature of the good, suppliers who are this case the physician, act as the patient agent for the following reason; health care is difficult to store and there is no substitutability across broad categories of consumers, different patients requires different treatments which mean that the commodity is customized. The problem or snag with this agency relationship unlike the usual commodity the market is that - all patients' demand for medical services comes from the suppliers meaning that the consumer cannot benefit from the efficiency of competitive markets. This then bring the problem of entrusting fully the welfare maximization of the consumer to

the hands of the agent who is more likely to be interested at maximizing his utility function more than that of the patient. Given that the provider offers both information and the service, there is likelihood of conflicting interest.

d. **Indivisibility feature of the health good**. Unlike other goods and services health goods are not stored in small fraction. It either you purchase a health care good in totality if you want to achieve it benefit or you do not. Unlike other goods whose storage are affected by perishability, health goods are consumed instantaneously with the need. Though one can purchase a first AID kit for the unanticipated future occurrence of health need, this is unlike other goods which consumers are certain to use in future. This Indivisibility assumption could lead leads to monopoly tendencies of the supply by the physician.

e. **Presence of insurance in health care market**. This make the commodity health unique as compared to many other commodities in the market. The presence of health insurance a market complicate the perfect competitiveness of the health care market. In a instances where patients have a health insurance, the purchase of health commodity is paid by the insurance company and in case of the co-payment the patient does not pay the full cost of treatment by assumption. Again, such separation from the direct payment must weaken some of the price effects that might be expected in standard economic analysis. Insurance changes the demand for care and it might as well change the incentive of the provider.

f. **The role of externality in health care**. Unlike standard commodities, health has an externality angle. Poor health of another person often evoke a feeling of concern that distinguishes health care from many other goods and services. This bring the adage that people need to have an amount of health they need regardless of their ability to afford. Consumption of health good causes positive externality to members of household as the sick person recovers and non-consumption of the same coupled with health state deterioration causes negative externality to the same household. Hence healthcare as is a unique commodity in its own sense.

Healthcare as a public good

What is a public good?

A pure public good is one that no one can be prevented from consuming and that it can be consumed by one person without depleting it for another. The marginal cost of providing the good to another person is zero. The most common definition of a public goods stress two attributes of such goods; non-exclusivity and non-rivalness.

Non-exclusivity:- a good is exclusive if it is relatively easy to exclude individuals from benefiting from the good once it is produced. A good is nonexclusive if it is impossible or very costly to exclude individuals from benefiting from the good. Again, public goods can be contrasted to excludable private consumption goods (such as food, cars, telephone, and payable TV stations) for which exclusion is a simple matter.

National defense is the standard example. Once defense system is established, everyone in a country benefits from it whether or not they pay for it. A good example of such good in health care could be mosquito to control or inoculation against a disease programmes. In these cases, once the programs are implemented, no one in the community can be excluded from those benefit whether or not he or she pays for them.

Non-rivalry: a non-rivalry good is one for which additional units can be consumed at zero marginal social cost. For most goods, consumption of additional amount involves some marginal cost of production. Consumption of one bread by someone for example requires that various resources be devoted to its production. For certain good however this is not the case. For example, having one more car pass through a bridge or road require additional resource use and does reduce consumption elsewhere.

Unfortunately, in the health care market this feature of non-rivalry really occurs. When we consider the case of mosquito control for example by the use of impregnated mosquito net, distribution of mosquito nets will imply depletion of the next household which will entail production of extra nets for the next person. On the same breath some

health programmes satisfy this feature. For example, if the malaria program was to spray breeding cites for the mosquitoes in a certain village, the benefit will be nonrival to households in that village because being in that village does not require another spraying done.

Assignment:
   a. Discuss the rational of considering some healthcare goods or service as public goods.
   b. Discuss using appropriate example different categories of health care goods or services that satisfy either one or both feature of a public good in your country.

Externalities and Market Failures in Market for Health care

The allocation of health goods frequently deviates from the principles of a market economy or from the various market structures. Moreover, the price mechanism is not permitted to coordinate choices in health care by signaling varying degree of scarcity. The deviation from market allocation are generally justified by claiming that health goods presents particular features rendering them different from other goods. These features are said to cause a *market failure* that result in an equilibrium that does not correspond to a *Pareto-optimal* allocation in non-regulated markets.

Most consumer goods provide private benefits and little or no externalities. Because demand will represent only the private benefits, it will understate society's benefit or cost and give false or inadequate signals to the market for not taking care of existing *externalities*. This causes inefficiencies which is referred as market failures
Let's now turn to what externality and market failure entails.

What are externalities in the health markets?

An externality is defined as an uncompensated direct effect of the production or consumption of a good or service on persons other than the producers or consumers. Externality occur when an action undertaken by individual (firm) has secondary effects

on others, which may be favorable or unfavorable. Individual consumers and producers in a competitive market do not normally take into consideration the external benefits or costs imposed on others as a result of their decisions.

Let us have an illustration on externality issue;

Consider two firms—one producing well X and the other producing good Y, where each uses a single input, labor. The production of good Y is said to have an external effect on the production of X if the output of X depends not only on the amount of labor chosen by the X-entrepreneur but also on the level at which the production of Y is carried on. The production function of X can be written as

$X = f(L_x ; Y)$,

Where

L_x - denotes the amount of labor devoted to good X

Y - is an effect over which the X-entrepreneur has no control

Assume firm Y pollute a river in it productive process while assuming that firm X is downstream from Y. The output of firm X may depend not only on the level of inputs it uses itself but also on the amount of pollutant coming from firm Y. In this case the output of firm Y would have a negative marginal physical productivity $dx/dy < 0$ increase in the Y causes a decrease in X meaning a negative externality. This formulation can also be adoptable when X output is benefiting from Y productive process.

To understand externalities, we need to understand how prices and cost are set in a complete markets. Markets work best when all costs and benefits of a product are taken into account in the setting of it is price in this case the market is said to be complete. The assumption is that all relevant information about the product is incorporated.

But this is not always the case as spelt under four outcomes

1. Consumption of a product could damage a third party: in case of a smoker, his enjoyment of puffing within an environment where people are cause contribute to their parting in his or her smoking through passive smoking. This is classic case of

a situation when a non-smoker's health is affected by the smoker's behavior. This is classic example of negative externality.

2. When consumption of a product benefits someone else: a good example is when a person getting vaccinated against an infectious disease protects the vaccinated and protect further transmission to others or alternatively when one purchases a vaccination for influenza. This good entails a private benefit, the purchaser will less likely suffer from the influenza. In additional, there is an external benefit to others because the purchaser will be less likely to infect others.

3. When product is produced and it damages a third party through air pollution which affects the environments, Industrial pollution has exacerbated serious health problem to the general population. Its effect on climate change has contributed to greatly to health hazard on both fauna and flora of this world. This is a very serious case of negative externality brought about by production process.

4. When product is produced and benefits a third party. Positive externality could also arise in production scenario when for example one decides to plant trees for either environmental purpose or for business purpose. The trees will also absorb carbon dioxide and provide shelter for animal species which is an added benefit that also improves the health of the population. This will not have been accounted for during the decision to plant trees.

These important attributes are not reflected in the price of the product-they are thus called externalities, because they have not been factored in the initial decision/objective. Hence externality will affect the true price of the good. If the consumption of a good h by individual I has a direct effects on the utility of individual j. giving rise to an externality, then the market mechanism does not lead to pareto-optimal allocation in a situation of perfect competition. The reason is that in this case , consumer optimally buys a quantity of goods h such that his or her own marginal utility derived from the last unit of the good is equal to its price. In doing this, she or he does not take the potential effects on j's marginal utility into account.

What is Pareto- optimality?

Pareto-optimality/efficiency- situation in which it is impossible to improve the welfare level of one party without hurting the welfare level of another party. Situations in which the level of welfare of one or more parties can be improved without hurting any other party are Pareto improvements.

Positive externalities are associated with under-consumption of goods while negative externalities are associated with over-consumption in a market economy relative to the Pareto-optimal solution. ***Why???(Underestimation and overestimation assumption)****. Negative externality is considered as an overconsumption because the harm of the good or service is underestimated which means it is causing more harm than the original one. For example, when one smoke the harm of smoking is extended to more than one person implying the consumption was is more than the initial. The implication is that we need to lower the consumption of such good. On the other hand, positive externality is assume that the benefit is only one again on the producer or consumer but in reality it is being passed to others. Hence while accounting for these benefits we need to add their marginal effect to the third party meaning we had underestimated the utility of the health good for instance.*

In the case of health goods, positive externalities are more relevant. See the Figure 4.6 In the Figure 4.6 MPB represent Marginal private benefit that does not consider external benefit from other. MEB represent marginal external benefit. MSB represents the marginal social benefit, which incorporate the MPC and MEB. MPC represents the Marginal Private cost. ***What does it show?*** It show that ***output before consideration of external benefit or positive externality is lower than after consideration of the benefit i.e. Q₀< Q₁***

Figure 4.6

With these illustrations, one can conclude two things:

a) It affects efficiency assumption in that it only when all the marginal benefits (MSB) equal all the marginal cost (MSC) that the optimal level of output is determined. Under the current condition this does not hold.

b) It effects on redistribution; here some persons receive an external benefit for which they do not compensate those proving it, or a cost is imposed on them for which they are not reimbursed.

Figure 4.6: Externality in production and consumption
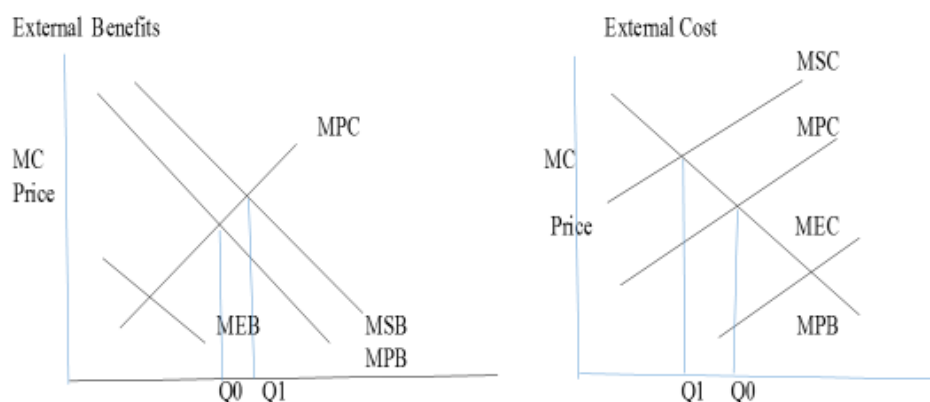
The consumption of health goods *h* by individual *i* can either

a) Directly improve the health of individual j (these are called Physical externalities)

b) Simply lead to increased satisfaction for individual j (these are called Psychological externalities)

Physical externalities

This directly improves the health of individual j. In a health good scenario, a classic example is in the context of communicable diseases. If individual i seeks treatment or tries to prevent that she or he does not contract such a disease (say COVID19) the

116

probability of individual j getting infected is reduced. In a two –person case involving I and j only (e.g., Vaccination),a Pareto-optimality could be achieved by j voluntarily subsidizing the vaccination costs of i. In reality externalities are spread over many people, giving rise to their collective good property which has got its own complications, A 'collective good' also called public good (already discussed earlier) is characterized by non-rivalry in consumption.

Here, if a consumer *i* obtains a unit of good *h, then another consumer j can also use the same unit without diminishing i 's* enjoyment of the good. The condition is satisfied in an ideal way in the case of the external benefit created by the vaccination (good h) of individual *i, since this affects several other individuals, even without a diminishing effect if their number increases.*

Again, another notable attribute of the collective good is non-exclusion. Nobody can be barred from using them even though she or he has not contributed to their provision. In a perfect competitive market, these goods will be under-provided. This because individual consumer has no incentive to contribute to the financing of such a good (for example in the case of COVID19, participate in financing production of masks for individual i), being able to benefit from the advantages of reduced risk of infection while other takes of the financing.

This argument suggests that the programmes like compulsory vaccinations, wearing of mask by the general public to prevent communicable diseases should be undertaken by government through use general tax. Government can also offer subsidy on the production vaccines or mask.

But again, this does not call or justify for a blanket provision of all medical care by the government. Given that not all citizen are exposed to Infectious diseases it positive externality will not be realized as such. Again, once individual learn how to protect themselves against infectious disease the positive externality vanishes. Conversely, if

those who are themselves not at risk are forced to contribute to the cost of controlling an epidemic through tax payment, then public provision no longer leads to a Pareto improvement relative to the market solution.

Psychological externalities-

This type of externality is based on 'altruism' behavior of the individual or society. This behavior leads to increased satisfaction for individual j which is not observable. This is emphasized by the fact that the utility level of altruistic individual j not only depends on her or his own consumption, but also (again positively) on the consumption of individual *i.*

Positive externality may arise as long as the consumption of fellow citizen is perceived as being unbearably low. Externality may also arise when people are affected by the suffering of others mainly in health care environment(caring externality)-people feel upset or sad to see sick people who cannot afford treatment and who are suffering and dying in their midst. It also causes most people distressed when seeing others die from starvation or lack of medical care particularly if their suffering was not caused by their own actions.

The key questions is

　　a. Should these externalities be internalized through private or some public intervention?

　OR

　　b. Should they require subsidization or provision of certain free goods (e.g. medical treatment) or through money transfers

　OR

A combination of both a and b

In the absence of the government a society with wealthy member contributing to welfare will result to sub-optimal allocation or provision due to the fact altruism behavior cannot be relied 100%. Charity therefore has the properties of a collective good which provides

an argument for the government to organize support for the poor who cannot afford medical treatment.

Points remember

- Externalities can then relates to consumption and production. They can be positive or negative.
- Thus, one can define externality as an uncompensated direct effect of the production or consumption of goods on persons other than the producers or consumers

Let provide an illustration of health good in terms of externality

To demonstrate this externality and market failure we can think of a case of treatment of tuberculosis using Figure 4.8

What would happen if treatment of tuberculosis were provided privately and paid for entirely by patients?

Figure 4.8 Demand for TB drug

Demand for a tuberculosis drug
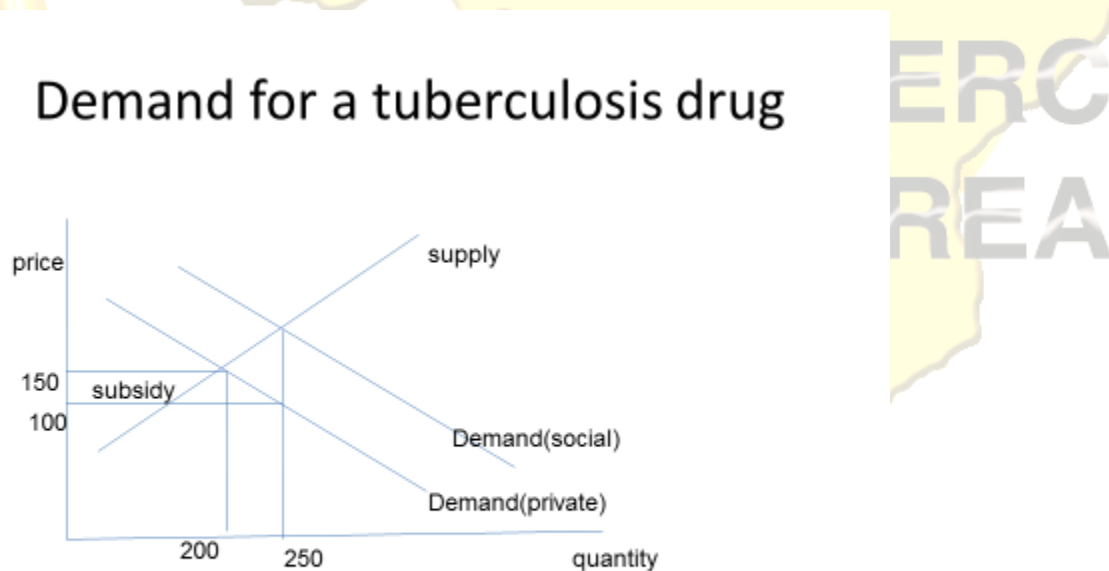
In the initial demand curve 200 courses of treatment would be demanded at 150 shillings per course. This reflect the private demand curve incorporating the value placed on treatment by individuals. What is omitted is the wider social benefit of slowing the spread of the disease. If social benefit could be measured and plotted, the real demand curve would be further to the right. *This means that a free market is producing less of tuberculosis treatment than socially optimal.*

One solution would be to provide a subsidy such as free drugs for tuberculosis treatment. Which would increase the demand to 250 courses which coincides with the socially optimal equilibrium. This is case where the community is consuming less of health good health with expected positive externality. The idea is to encourage consumption of more of such good because the higher the consumption the higher the benefit.

What would be equivalent of a negative externality?

Answer: Tax in case of alcohol and cigarettes. In order to discourage negative externality of alcohol and cigarettes, government in most countries of the world have used taxes instrument as a way of discouraging consumption of these goods which are known to have negative externality.

On the other hand, policy makers may be affected by the positive effects of good health on health productivity. To this end, health care may be seen as a merit good-a good which we would like people to consume more than they might if left alone. Government can distribute education materials urging people to have dental check-ups, eye- checks, blood pressure checks so that they seek early care hence reducing high cost related to curative care in future. This will save government money in future which can be used to improve health sectors in other ways. This can be viewed as public authorities trying to increase private consumption on health to help avoid future illness thus saving money for public purse.

Assignment:

Explain how transfers would be used to solve the problem of psychological externality in a health care setting. Provide a critic for both use of **in-kind and cash transfer** in this case.

Market Failures

Market failure occurs when the price mechanism fails to account for all of the costs and benefits necessary to provide and consume a good. It occurs due to inefficiency in the allocation of goods and services.

What are market failures in health care markets?

It is a situation in the market where the market provides less than the amount that would maximize social benefits. In an ideal situation, markets have an ability to produce goods efficiently, to channel resources to their most valued use and to maximize social welfare in general sense. Market failures is thus known to result in inefficient outcomes and to change distribution of resources among the population As noted earlier there are some integral characteristics of health care that affects this Pareto- optimality.

Causes of market failures

There are three broad causes of market failures. All the three have a bearing on health care provision as will be explained. They include

1   Structural deficiencies existing in the market
2   The product in question is a public good.
3   Externalities associated with the products

Let us look at each before we move to health specific market failures and solution to each.

Structural deficiencies

Structural deficiencies arises due to market imperfection or market power and information asymmetry or imperfection. Market control occurs when either the buyer or the seller possesses the power to determine the price of goods or services in a market. The power prevents the natural forces of demand and supply from setting the prices of goods in the market. On the supply side, the sellers may control the prices of goods and services if there are only a few large sellers oligopoly or a single large seller (monopoly). The sellers may collude to set higher prices to maximize their returns. The

sellers may also control the quantity of goods produced in the market and may collude to create scarcity and increase the prices of commodities. On the demand side, the buyers possess the power to control the prices of goods if the market only comprises a single large buyer (monopsony) or a few large buyers (oligopsony). If there is only a single or a handful of large buyers, the buyers may exercise their dominance by colluding to set the price at which they are willing to buy the products from the producers. The practice prevents the market from equating the supply of goods and services to their demand. If these occur government intervention could possibly result in greater efficiency than relying on the market alone. This is only a possibility due to the fact that inefficiencies entailed in a government activity might offset the potential gains. But it is important to note if we recognize that market may fail we may also recognize that government may also fail.

Market failure may also result from the lack of appropriate information among the buyers or sellers or what in the literature is called information asymmetry. This means that the price a good does not reflect all the benefits or opportunity cost of a good. The lack of information on the buyer's side may mean that the buyer may be willing to pay a higher or lower price for the product because they don't know its actual benefits. On the other hand, inadequate information on the seller's side may mean that they may be willing to accept a higher or lower price for the product than the actual opportunity cost of producing it.

The product in question is a public good

Public goods are goods that are consumed by a large number of the population, and their cost does not increase with the increase in the number of consumers. Public goods are both non-rivalrous as well as non-excludable as seen earlier. Public goods create market failures if a section of the population that consumes the goods fails to pay but continues using the good as actual payers. For example, police service is a public good that every citizen is entitled to enjoy, regardless of whether or not they pay taxes to the

government. You can also think of vaccination: vaccination is a private good with external benefit. The external benefit of this good entails certain public good aspects.

Note if pure public goods were to be provided by the private enterprise it will find most citizens choosing to be free riders. A free rider is a person who consumes public good but refuses to pay (e.g. security in a region provided by a community). This is due to the fact that a private enterprise has no power to tax neither does it have power to exclude the services of the public good to non-payers.

Externalities associated with the products

As seen earlier an externality refers to a cost or benefit resulting from a transaction that affects a third party that did not decide to be associated with the benefit or cost. To this end externality results in non-optimal amount of output being produced because an individual or firm consider only their own benefit and cost when making consumption or production decision. When cost or benefits are received by others as a result of own private decision the level of output produced in the market will be based upon either too small a level of benefits(i.e. positive external benefit) or too small a level of cost of production(i.e. negative externality or positive external cost).

Some example Market Failure In health and how to tackle them
Option-Good property of medical services.

A peculiarity of most medical services is that at individual level, their demand are not predicable. When it arises however it may often arise on an extreme urgency basis while capacities for providing medical services particularly in the hospital, cannot not be created at such a short time. This simply mean urgency is not consistent with ability to change provision within that urgent notice. This create a serious market failure in the provision of the medical services. To solve this kind of short fall in the provision of such services, hospital need to create some reserve capacity that should handle such uncertainty. The existence of such reserve constitute what is called option-good property, which indicate that the mere existence of the capacity yield utility to the consumer by knowing that uncertain need is taken care of in future. In order to create

an incentive for hospital to hold capacity on reserve, this must be paid for separately. Implying that the hospital's revenue must not depend on bed occupancy only.

The question is where does the hospital get such reserves?
From the Government? No and yes! Option good can be meet within the domain of private health insurance. The insurer guarantees the insured that a hospital bed will be available in case of need. Part of the premium revenue is then used to pay a hospital in the insured person's area of residence for holding reserve capacity.

Yes: If the good this time has a characteristic of a collective good property or a public good property say for example an epidemic or pandemic like COVID19. This financing of option good can be done from the taxes. The condition of non-rivalry is fulfilled, since one and the same hospital bed can provide several potential patients with the utility of secure access at the same time.

Another question: Will exclusion principle then hold here in case of pandemic like COVD19 where the bed capacity will get filled up even after option good property has been implemented through tax financing**?**

This is a hard policy question. The answer is two sided.
One, the non-exclusion criteria will be by passed in time of shortage by admitting only those who have concluded an appropriate health insurance.
Two, since in an emergency, the exclusion principle must presumably be waived for humanitarian concerns, it is justified to force all citizens to carry insurance for emergency access to a hospital. Such a mandate is not equivalent to financing emergency departments using tax money because competing health insurers can negotiate for their provision.

Lack of consumer autonomy/sovereignty
Another frequently cited reason for market failure in health care is consumer are not in a position to decide amount of medical care rationally due to two things:

a) Incapability of rational decision making

b) Excess discounting of future needs

Incapability of rational decision making

It argued that a person in state of unconsciousness, great pain, or mental illness admittedly cannot make a rational decision. Consumer sovereignty is challenged under this condition. There are doubts that individual under these conditions are capable of seeking treatment out treatment options that maximizes their utility taking into account the associated costs in a rational balancing process. While in the said state a person will not be able to make a rational decision. Close relative or guardian would be expected to make a rational decision on his/her behave.

Another problem for irrationality is if the illness or disease is life threatening while not limiting mental capacity. The Patients at this point are in weak position relative to the supplier of health services since they are prepared to any amount asked by the supplier. The purchase of health insurance would take remedy such patient's structural inferiority to the provider of care. But the downside of this is that patients have low interest in seeking out low-cost treatment alternative.

Apart from competitive market-oriented solution is there a solution to this consumer sovereignty?

Answer is yes

One:- before a life-threatening situation arises, individual can take care by selecting someone close to make the decision for them

Two:- Purchase health insurance which can ensure that they are not exploited by health providers.

Excess discounting of future needs

By nature, frequency and severity of illness increase with age, causing the likelihood of treatment costs that cannot be paid out of current income to increase as well. Figure 4.8 show an example of private per capita health in Kenya which mimic global average per

126

capita health expenditure trends. The trend shows as people age the health expenditure rises. This is what is called j-shape for health expenditure.

Figure 4.8 Example of Kenya private health consumption below

Kenya Private Health Consumption, 2005
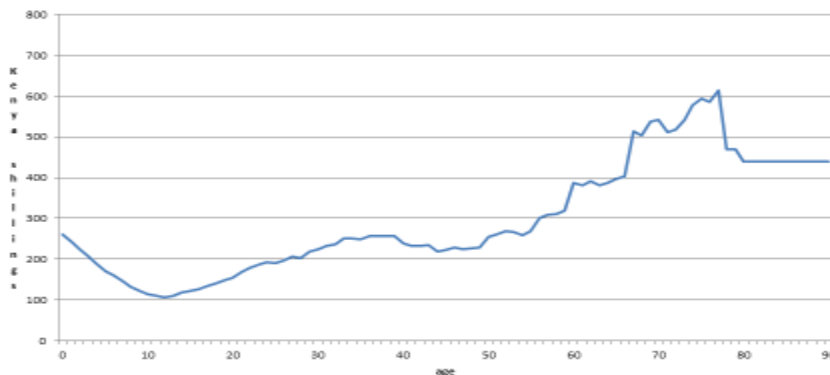
A rational individual will have to save for future health expenditure or will purchase a health insurance policy. But given that most people value future consumption less than present, however, May will only attach limited importance to this eventuality thus causing insufficient coverage for health care in future and at old age. This will lead to what is called - a systematic under provision of health goods.

The big debate is who should decide on how future health provision will be voted for? Is it the youth or the old?

This provision will depend on whether it is supported by the majority of voters.

Youth would prefer current consumption of goods in the market rather than uncertain future. On the other hand, the old, though have gone through negative experience of sickness, would perceive this program as not useful since they are not likely to benefit from it.

Assignment:

Discus methods that we can use to solve future market failure in health consumption of health brought about by excess discounting for future.

Lack of perfect information

Lack of perfect market transparency in the provision of health good causes market failure. Pareto optimality is not achieved due to this condition. Market transparency is a condition where potential buyers are informed about the qualities offered and prices quoted by all suppliers on the market. Perfect knowledge of products quality is not assured in cases where production and consumption occurs simultaneously ('uno actu principle'). Under uno actu condition you cannot compare the offers. Medical services fall under uno actu principle just like restaurants, artistic performance and service of hairdressers or barber. Three additional features can be identified which distinguishes health goods from most other services.

a) Lack of opportunity for sampling

b) Unclear cause effect relationship

c) Asymmetric information.

Lack of opportunity for sampling

Medical services unlike services such like restaurant and particularly those crucial to survival in the case of life-threatening illness are often consumed only once in a lifetime, causing a lack of experience for evaluation. Additionally, experience of others does not help since health problems are not comparable because treatment process is mainly patient- physician component. This interaction between patient and physician makes an objective evaluation of quality difficult.

Unclear cause effect relationship

Most of the time, the quality of a medical service cannot be judged correctly even after utilization, since the causality between treatment and change of health status may be affected by other biological processes such as the self -healing power of the body.

Asymmetric information

In the case of diagnostic services in particular, the good demanded is information. This makes it impossible for patient to evaluate the quality of service prior to its purchase

because this would require them to already be in possession of such desired information. The information advantage of the supplier give them a certain degree of power over the buyer as seen in the section on PID or SID. This makes medical services very unique in this respect.

Assignment

1.   What is the way forward for information imperfection in the health care market?

2.   Discuss how government and health insurance would-

a.) Help in solving market failure in health care market

b) Aggravate /worsen market failure in health care market

Role of Government in HealthCare.

From our discussion on market various one would be forgiven to assume that movement toward perfect competitive market is the solution to health care market. This appears to be the hypothesis that market should be relied on more in the health services sector. But knowing the characteristics of our health goods this will handily work. The alternative to relying on markets is varying degree of government involvement in the and financing and delivery of these products and services. Although nearly all health economist agree that government must play a role in the market for health, the nature and magnitude of such involvement is controversial.

The question of market failures where markets are problematic not only with respect to providing services equitably, but also efficiently begs the key policy question; would government do any better. There those who feel that the inevitable drawbacks of government in terms inefficiency could do more harm than the shortcoming of the competitive markets. Whichever way we argue, government intervention in the health system is paramount as provided in the following discussion.

Let's now turn to the role of government in especially battling disease in the country

a.   Government intervention on the economic analysis is seen as a result of market failure or when issues of equity arise. Traditional economic theory provides two main sets of reasons for government involvement in a market; to overcome market failures that results inefficient outcomes and to change the distribution of resources among the population. An area of concern when it come to market various is the issue of externalities that is attached with the health goods. This can be discussed through how government try to battle with the diseases. For instance, where immunity to the disease can be obtained through vaccinations or other means, and where disease is contagious economist commonly recommend that vaccinations be encouraged either through provision of subsidies. The reasoning is that private pay vaccination will tend to underplay the external benefit of immunization i.e. the immunization gain not only benefit the vaccinated but also other. By not becoming disease carriers, they reduce the risk to others. This is what is called externality rationale for collective action.

Note. *In order to justify the intervention of government in market for health goods one need to show that at least one of the requirements cited in the market failure is in place, that is*,

➢   Health goods have the characteristic of public goods.
➢   The consumption of health goods gives rise to externality.
➢   The criteria of a perfectly competitive market, specifically market transparency and consumer sovereignty are not meet.

b.   Another concern for government intervention is to achieve a kind of optimality in disease control. Poor consumer information is one reason that markets fail. In a competitive market information is expected to be adequate and symmetrical to the market participant. Infectious diseases commonly fail to fit this description. Risk faced by population due to spread of an infectious disease may not be fully known. Thus, given that Government can invest in information search, it is worth giving a room for intervention to take care of any disease spread. But proponents of markets often believe this is not a justification for government intervention because if anything this will be ineffective due to the fact information cannot be made

understandable to or policy most people or the people may not use information the way considered desirable by experts or policymakers.

c.   Another reasoning for government participation is on Equity issue. Equity may be violated, for example should segments of the population suffer the disease more than others, it might appear like the disease has been forced on them through some public actions. Hence it is the role of government to ensure that quality health is accessible to all their citizen. This is done through universal health coverage in most countries. Universal health coverage help in the following ways; it improve the health and productivity of the population by making health services financially accessible, obviate the need to provide for a large array of safety –net facilities for uninsured sick people who cannot afford care , reduces the problem of adverse selection into health insurance plan and enhance fairness in society. With this notion of universal coverage, it attests to the prevailing notion of fairness where people should not be penalized for circumstances such as their sociodemographic background or their current state of health.

Another role of government in health intervention on equity issue is in the transfer programs. Cash transfer programs are intended to meet society equity concerns by redistributing income, with recipients free to spend their income in any way they want. Government also provides in-kind transfers with the main purpose being to increase recipient consumption of a specific good. In-kind has been used in many health programs especially where a certain population more so the poor is not able to access adequate health goods.

d.   Another argument for Government for Government inclusion is its concerns or macroeconomic consequences of the diseases. A disease like AIDs left alone could result to heavy consequences or the productive capacity of the economy. This will affect growth of the GDP and thus further depletion of Government revenue in the form of expected Tax Revenue. Government can adopt a variety of policies and instruments to influence the allocation of resources or distribution of income to enhance macroeconomic stability as much as having a twin objective of

ensuring equity. The policies can center on commodity taxes and subsidies-increasing taxes to goods with negative externality in order to boost revenue going to health sector, provide subsides to the producer of goods of goods that are health improving. The policy can also center on provision of public goods like public provision of healthcare in which the funding mechanism could have large impacts on the resources allocated to health and on the possible re-distributions as from rich to poor and young to old.

e. Apart from taxes and subsidies, government also influences the allocation of resources by establishing rules and regulations. Government can prohibit certain goods or activities entirely, such as prohibition of production of illicit drugs and alcohol which have health hazard implication.in most instance governments regulate the form and terms under which goods are produced or consumed. Regulations in health care markets can take many forms, for instance; licensure, laws, mandates and regulation of prices, quality, and quantity of services as will be seen later in this section.

f. Another argument for Government intervention is its intermediary role between charitable organization and health programs. A health program like reduction of AIDS is funded majorly from donations coming out of the county. Government control how WHO's and other community-based organizations undertake their role in these health programs. Government also takes role in the management of funding from global charitable organization and individual on health programs.

Though government intervention is seen as Economic justifiable, those concerned with economic and social behaviors argues otherwise. For example they should ask government encourage and provide AIDs testing? The argument is that those testing negative may increase their risky behavior or those testing positive may not care about infecting others. Thus to understand, predict and control the epidemic, it may require the understanding and manipulative of prices and benefits that those involved face when making their choices.

In sum, hardly any economic activity or sector is free of government intervention. Health is no exception especially with the known market failures. Intervention comes mainly through three activities; provision of goods and services, redistribution, and regulation. Government have pursued each of these activities in the health economy. Notable is the provision of health care and of health insurance is a major mean means used to redistribute income to lower-income groups from higher income-groups

Market for Health workforce

Health economy constitute a high share of gross domestic product of all the countries. Some countries spend close to 20% of their GDP on health. The labor force requirement is huge for the health sector, example 2003, 10% of the US labor force was in health economy. This data did not include millions of other workers in pharmaceutical and health insurance industries. Like any other goods and services, the production of health services requires both labor and capital.

Let us now look at how demand and supply of health work force is determined in a competitive market before turning to an example of physician market
Demand for and supply for health work force

Demand for a factor of production either labor or capital is derived from demand for health. We demand healthcare provider because we demand healthcare, and in turn we demand healthcare because we demand health. The demand of any type of healthcare labor depends in part on substitution possibilities as much as it depends on it the price of output. See the substitution possibilities in **Figure 4.9.** The first figure shows a case where factor substation is possible where labour and capital are good substitute as one

move along the isoquant curve. Lower prices of labour will lead to a relatively larger substitution of capital for labour and vice versa. The second Figure 4.9 shows technology in which labor and capital are not good substitutes, they must be used in fixed proportion to one another. In health care one can think of specialized surgeries which require specific ratio of labour to capital with little substitution available. Thus, the degree to which substitution among inputs is possible , either between health care labour and capital , or among different types of health care labor, is key issue in manpower planning and in determining the efficiency of production exhibited by health care firms.

Figure 4.9 production with varying rates of substitution
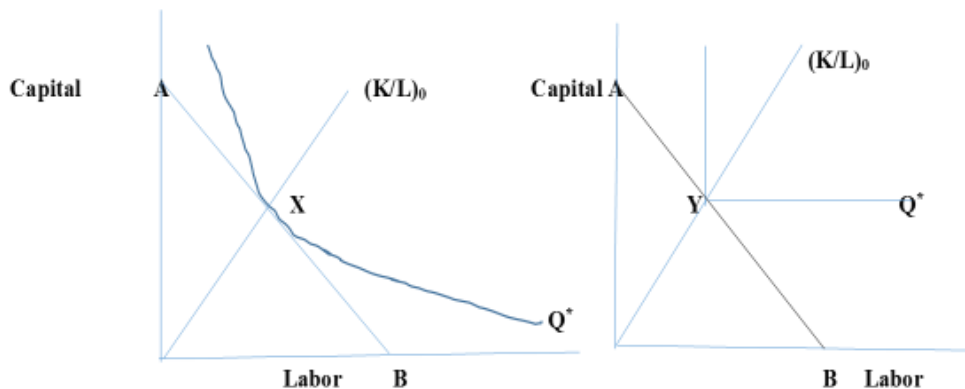
The supply of labor on the other hand tends to slope upwards implying that the higher the wage rate in given market, the more workers or labour time will be forthcoming. This due to two reasons; first those workers currently employed may choose to work more hours if higher wages are offered, second, similar workers from elsewhere may be attracted into the market due higher wages. Figure 10 show labour supply for laboratory technicians**.**

We have two equilibrium points' w1 and w2. At W1 the number of technician is will and available to work at wage rate is L1 which is also the market equilibrium point where demand for labour is equal to supply. If there is a demand increasing event that shift demand from D1 to D2, the supply of will increase along the supply curve to L2 which is the equilibrium point of this laboratory technician labor market. The events that could contribute to increase in labor demand firms desire to substitute laboratory technician for other laborers, increased productivity of technicians or increases in the prices of lab output.

Changes in supply can take place relatively quickly for those health care occupations requiring minimal education or training. For physicians specialists and other requiring highest level of education and training, the number of new professionals is determined by the admission decisions of medical schools as well as their personal decision that they made many years back. Through education and training medical students make investment in their human capital. This is based mainly on the expected rate of return in this field after successful completion and employment in the health sector. Information regarding rate of return to a medical education is useful to the policy analyst. If analyst for example discover that the rate of return to medical education is on average increasing then they can conclude that physician are becoming increasingly scare.

Figure 4.10 Market demand and supply for laboratory technician
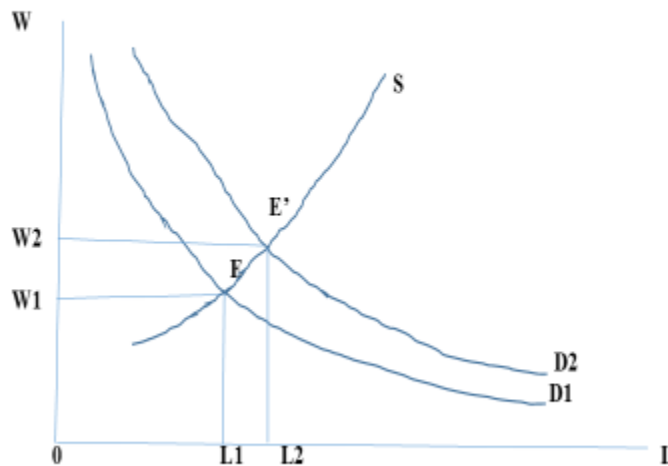
Let us now have a look at the market for physician as a case of health workforce.

Market for Physicians

The patient demand for medical care is the demand for treatment which calls for the demand for physician time. Hence the factor affecting demand for medical care are tied to the demand for physician time. Hence, demand for physicians services is determined by both economic and non-economic that are linked to the patient demand medical care-

Economic factors such as:

Prices that the physician is charges. Knowledge of the price can be signaled by patient experience or information set from patient network. Prices of complementary and substitute services will also affect the demand of the physician. If patient have close substitute like having a pharmacy that purport to have physician. This will mean that patients will perceive the services of such alternative to be much cheaper compared to a professional physician. The price of complementary services will include such cost like

137

laboratory tests, travel cost to the physician, prices of medicines cost of consultations if charged separately from the medical cost and nursing cost. Another complementary issue in the demand set is time cost (waiting time) that are involved in the purchase and use of physician services. For very specialized medical services like neuron surgeon waiting time is normally is quite long to an extent that patient feel fatigued queuing for many hours.

Income of the patient is key in the determination of the demand for physician time. Income not only affect whether the individual demand treatment but also on where and who to seek treatment from. If individuals have a Permanent the demand for physician time is more stable because they cannot only seek curative services but preventive services as well overtime.

Insurance cover is another key determinant of demand for physician time. Individuals will not be constrained by their income when choosing which physician to visit neither will they be constrained by treatment cost though this will depend on the comprehensiveness of the f health insurance plan

Non-economic factors such as:

There are many factors that affects the demand for physician which are outside economic consideration of the patient. Let us look at some of them

Trust the patient or the family has on the efficacy of medical treatment shape the consumer desire on whether to seek the services of the physician or not. Consumers can bypass several physician on the basis of trust they on one particular physician.
Distance to the physician though it has an economic perspective is considered a key determinant on physician demand. Theoretically the further the physician is to the patient the higher the likelihood of reduced physician visit.

Quality and process factor is key to determining the demand of the physician time. Quality of the service is perceived as good if the patient is able to get the right treatment in terms perception of the patient. In some instance quality could signaled by prices which patient perceive to mean high quality. Process is the way the patient

Gender of the physician and the patient has bearing on the demand for physician time. This I more a cultural issue than a reality. In some society men will want to be attended by male physician for they feel like their privacy is interfered with when taking of their health condition to female physician. Again, studies have shown that women seek medical care more than men. This could imply in area where we have higher number of women physician demand will tend to be high.

✓ Education is also believed to affect the demand for physician time. Education is in two fronts. We have the education of the provider and the education of the consumer. Providers who are believed to acquired specialized training like posing a master degree in an area tend to be in high demand than their counterparts. The education of the consumer may enable him or her to recognize the early symptoms of illness and thus demand physician time more than one lower education.

Supply of physician services is affected by among other factors:

➢ Price received for physicians (to depend on input productivity and prices)

➢ Number of physicians which will depend on among other things:

▪ Medical school capacity

▪ immigration

▪ Brain drains

▪ Death and retirement of physicians

➢ The amount of hour worked by physician

➢ The capital and equipment available

On the supply side of the physicians, the number of physicians combined with other inputs determines the available supply of services. Physician's services maybe provided using different combinations of physicians and other health manpower (auxiliary-nurses, lab technicians etc.). Greater degree of delegation lenders the physician to be more productive in the market for health care. Delegation will affect quantity, type of physician services and cost of medical services

In the short run all the inputs except physician can be varied. What is notable is that physician hours can be varied depending on the labor-leisure trade-off. If the price of physician hour increases it implies earning of more income with in fewer hours. This can affect supply depending on the physician objective. If the physician is a target work then high prices will lead to a backward bending supply curves this is because the physician will achieve objective within a shorter period a thus affecting supply in the short run. Even non-target worker could have a higher preference for leisure too hence achieving their objective within fewer hours again undermining supply in the short run.

The long run period of the physicians relate to a period when the number of the physicians can be increased. This period can be defined under changes in the number of physicians due to death and retirement, immigration of foreign trained medical graduate, brain drain to other countries etc. Medical training will have an influence on the long run supply of physicians. Demand for medical training will depend on the return to such education as seen earlier. It is assumed that the physician services are produced at minimum cost but the price at which they are sold might greatly exceed the cost. If this happens then then fewer physician services will be purchased than if the prices were lower. To this extent a good measure of how well a market for physician performs is the relationship of prices to costs in the market.

This cost-price relationship is often difficult to measure, hence it be inferred through the method by which physicians determines price in the market. Under relatively competitive markets, prices will approximate the costs. In monopolistic markets, the

sellers price may greatly exceed the cost of providing that services. ***Why? This is explained under the structure of physician market and* Economies of scale in medical practice**

The structure of physician service market help to understand how physician determine their output, prices and profits as well highlighting different market structures that physician can operate in. This structure is determined by barriers to entry and economies of scale to be gained by practicing in medical groups.

These barriers to entry include:

o   Licensing

o   Graduate from an approved medical schools

o   Restricted number of spaces to medical schools

Though the barriers exist, given the large number of practicing physicians, it is still possible that market for physicians could be competitive.

What about the economies of scale characterization that will influence the supply of physician time?

Economies of scale will be determined by how physicians choose to operate. A physician can operate as solo or as a group practice.

Solo mean that he operating individually as private and independent person. Solo practice in the market for physician services have been are heavily competitive markets with exceptions of limit of physician advertising and insurance company ability to negotiate discounted prices among physicians which then make them be in a position to differentiate prices. With some law imposed on restricting advertising and price competition, the structure of physician market has kept on changing.

This has evidenced physicians operating as group practice which is explained by the following factors:

•   Physicians could lower their cost if they were to take advantage of any economies of scale in group practice. For instance, sharing the cost of inputs- hence each

physician has the incentive to increase the use of input if the benefit exceed the cost of inputs.

- With increase in size the physicians has no incentives to monitor the behavior of others unlike in the case of individual physician.

- It is easy to reduce inefficiency in group practice since decision making under group practice is done by a physician committee and management rather than by individual. The only shortcoming to this is unless compensation in a group is done based on productivity there is a chance that group members will have less incentives to be productive. Again, in case of dispute especially regarding payment incentives to maintain productivity, the group practice are likely to dissolve.

Empirical result by Reinhardt (1972) showed that physicians in group practice generate between 4.5 to 5.1 percent more patient's visits than do those in solo practice. Another empirical finding by Frech and Ginsburg (1974) found large practice (group) were more efficient than solo practice concluding that there was presence of Economies of scale in the production of physician services.

Apart from cost there are other several reasons that explain the growth of medical groups:

1. Lack of patient information- presence of informational economies of scale. Being a member of a medical group convey information to the patient about the quality of service

2. By joining a medical group, the group reputation is immediately transferred to the new physician. This explained by the fact when a new physician enters the market he or she disadvantaged in the sense that it take time to build reputation among patients, but by joining the group the new physician is able to receive patients from other busier physician from the group.

3. The reputation of the group is more important to the patient for those physicians services that are less frequently and that are more difficult to evaluate. Family practice is less effective in capturing such patient. Finding a physician involves

greater search costs on the part of the patient as well as for insurers or employer contracting for physician services.

4. The desire for the physician to reduce uncertainty and share risk. Physician in Solo practice larger are likely to experience greater variation in workload and income than member of a group who share workload. The larger the size of the medical group, the smaller the uncertainty (variation) over one's annual income. But it worth mentioning if the size of the group becomes very large, productivity incentives will tend to reduce as the group members receives less of their marginal revenue.

Large medical groups have a greater market power than solo physicians or small medical group. There is less competition, market is more concentrated and they are likely to be able to increase prices. They are also able to market their services directly to insurers' employers and hospital. It is less costly for insurances or employer to contract with a single large group than with a with many individual physicians.

The main advantage of group practice is the economies of scale which includes the sharing managerial expertise, technical expertise, claim processing and marketing, informational economies of scale, reduced uncertainty and risk to physicians and increased market power. But note, the primary disadvantage of group practice lies in the reduced efficiency incentives as the size of the group increases.

In summary, the large number of practicing physicians and relatively small medical groups suggest a market structure that could be highly competitive. But, high level of insurance coverage for physicians, together with the lack of consumer information has resulted to market imperfection under physician services thus limiting the market to operate competitively.

Let now turns to how physician market operate
**Monopolistic competition**

The physician service market has attribute of monopolistic market since because there is existence of large number of competitor, the services are somewhat differentiated services and Faces a downward sloping demand curve. Neither physicians nor their services are homogenous as explained by different aspects e.g., medical school attended, age, sex, communication skills location of their services, waiting time of the clients and innate of the physician among others. Lack of information by patient on other physicians make the physician to be a less substitute to others Due to search cost of gathering information on different physicians some patient will experience high prices among the current physicians. Notable is that if prices in the physician service market appears identical, it is more of a pricing collusion and price fixing rather than competition. In contrast to a perfectly competitive market, a physician who raises prices would not loss all their patients. Since physicians differ and patient have limited information on other physicians, each physician faces a downward sloping demand curve shown in Figure 4.11.

The figure show DD is the overall market demand for physician's services, dd-demand curve is demand curve faced the individual physician. Note that dd has more price elastic slope while DD is somewhat inelastic, The dd curve indicates that if the individual physicians raised price and others did not, the physician would receive an increase in market share. Again, if physician raised price and others did not, they would lose market share. The price elasticity of dd exist because there are better substitute for each physician's services. The market Demand curve suggest that there no less substitute for all the physician services. Hence empirical finding that the demand for physician services is price inelastic would hold for overall market demand.

Figure 4.11 Imperfection of Physician market

Imperfection of physician market….
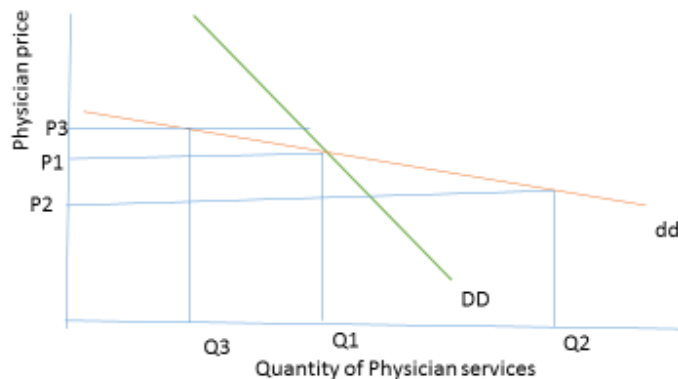
Monopoly Model of Physician pricing:

Let us now turn to the monopoly model of the physician

The model explains both variations and increases in physician fees when there is a decrease in the price elasticity of demand for physician services. The Figure 4.12 summarizes the model.

The model shows Imperfection of physician market where initial demand and marginal revenue curves are D1 and MR1.The physician MC is assumed to be constant. Resulting price under D1 is P1. This assumes an increase in insurance coverage which makes demand curve for physician visit to be fairly inelastic shifting from D1 to D2. The marginal revenue curve also shift rightwards. Intersection of a steeper MR2 with MC results in a higher price P2 on the demand curve. The economic theory implication is that variations in physician fees, unrelated to differences in costs, could be due to differences in price elasticity of demand. Which means that a possible explanation for the rapid increase in physician fees has been due to change in the price elasticity of demand for physician services. With insurance coverage patients have become less

145

sensitive to prices. This has tended to lowers price elasticity of demand and thus precipitating increases in prices by the physicians.

Gaynor and Polacheck (1990) estimated the degree of patient and physician ignorance in the market of physician services. The difference between the physician's fee paid by the patient and the lowest available fee charged by the physicians is assumed to be the measure of patient ignorance. Hey found that the degree of patient ignorance varies by type of visit, price dispersion is greater for services purchased less frequently, associated with greater severity of illness and those which are more heavily insured. They also found that patient ignorance is greater for those physicians whose practice relies more on physician referrals; referrals are a substitute for patient search and since patient search less, they have to pay higher fees.

Figure 4.12 Monopoly model of Physicians

Monopoly model of physician pricing
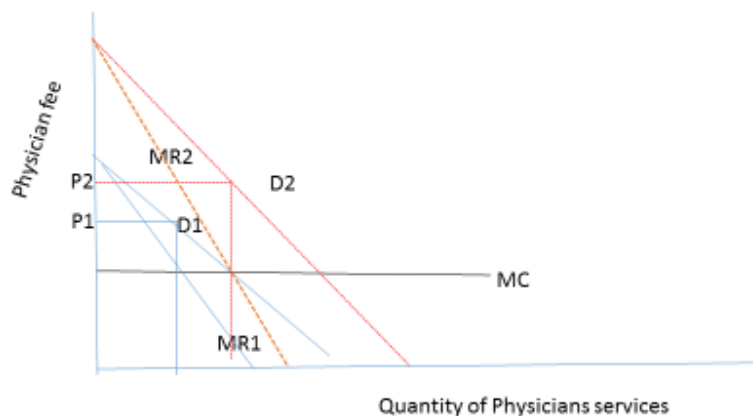
Markets for Pharmaceuticals

Pharmaceutical industry and prescription of drugs is at the center of current health economy. Today, drugs are used to treat many diseases and conditions, for example;

146

- Chemotherapy for cancer
- Steroids for skin diseases
- Psychotropic drugs for mental health
- Beta-blockers for heart disease
- Clot bluster for stroke
- Protease inhibitors for AIDs etc.

The good thing is some of these drugs are alternative to surgical procedures. Drug therapy has have supplemented nutrition as a way of preserving health. It is unthinkable to imagine of a world without anti-biotics, vaccines for smallpox and insulin for Diabetes. Pharmaceutical industries are among the largest and most profitable business in the world. It has variously reported that the cost of drugs has overtaken the cost of physician as reported by third party player (insurances). With prescription of drugs rising more rapidly, the industry is a convenient target for budgetary challenges facing patient.

This is complicated further by emergence of new and expensive drugs which always undergo vigorous direct marketing. The health burden due to drug spending is falling more to those patients with chronic and other conditions requiring frequent usage of drugs.

Pharmacoeconomics-a branch dealing with cost benefit, cost effectiveness and cost utility analyses, have provided some policy orientation aimed at mitigating the high spending on drugs that results to under provision of health care products.

Such policies include:

(i) Expanding insurance products to include outpatient prescription of drugs

(ii) Regulation of drugs prices as well permit the re-importation of drugs from other countries

(iii) Increased emphasis on generic products to help lower cost.

(iv) Development of drugs formularies so that patients obtain reimbursement for any drugs that is on the list.

Pharmaceutical industry is among the largest and most profitable business in developed economy. Their completive edge is driven by huge profits that is geared towards research and development thus making them even more competitive due adoption of new technology and development of more generic drugs. Pharmaceutical industries spend a huge part of their cost on advertisement and information – in 1980, German pharmaceutical industry spent close to 26 percent of total cost on advertisement.

The level of competition of competition holds the key to firm and industry behavior. Economist measure competition within the standard industrial classification (SIC) developed by the U.S. Department of commerce to categorize firms into meaningful industry and product group.

African Context on Pharmaceutical industry

A growth scenario is the one expected on African pharmaceutical market. The predicted growth in the global pharmaceutical market between 2012 and 2021 is estimated at 26% (according to PROPARCO's SP & D publication), but the African drug market could increase by 200% according to a realistic assumption, or by 300% according to an optimistic hypothesis within this same period

Based on African Pharmaceuticals Market Report of Frost and Sullivan of 2016, the pharmaceuticals market in Africa is expected to reach a business opportunity of $45 billion in 2020.

Causes of Growth in Africa Pharmaceutical

This rapid growth could be dependent of

- Convergence of changing economic profiles,
- Rapid urbanization,
- Increased healthcare spending and investment,
- Increasing incidence of chronic lifestyle diseases.
- The expansion of healthcare capacity which according to the World Bank Africa acquired 70,000 new hospital beds, 16,000 doctors, and 60,000 nurses, adding

up to a total of 1,050,000 beds, 498,370 doctors, and 1,250,000 nurses between 2005 and 2012.

African pharmaceutical market scenario

The tropical climate of Africa makes the continent the **largest reservoir of infectious diseases**, particularly malaria, tuberculosis (TB), and acquired immune deficiency syndrome (AIDS), besides frequent outbreaks of polio, meningitis, cholera, pandemic influenza, yellow fever, measles, hepatitis, and tetanus.

With the increasing adoption of Western lifestyle in Africa, there has been a **paradigm shift in the burden of illness towards non-communicable diseases** (NCDs), driving the demand for chronic prescription drugs. Based on World Health Organization prediction, the proportional contribution of NCDs to the healthcare burden in Africa will rise by 21% through 2030.

In the meanwhile, population will continue to suffer from infectious and parasitic illness, but lifestyle diseases such as cardiovascular diseases, diabetes, and cancer will witness high growth rates throughout the forecast period. Limited affordability of governments and the general population for healthcare and pharmaceuticals and a high reliance on donor funding will be the major market. Half of Africans do not have access to essential drugs. With the provision of the right drugs to treat respiratory infections, diarrheal diseases and malaria, around 10 million lives would have been saved in 2015. Every 30 seconds one African child is killed due to malaria**.** The 10 largest pharmacy retail chains in Nigeria, Kenya and Ghana manage 186 outlets for a population of almost a quarter of a billion people. South African companies hold the number one position for overall market share, for Pharmaceuticals only and for over-the-counter medication only (Lucchini, 2018).

African continent overall has roughly 375 drug makers, most in North Africa, to serve a population of around 1.3 billion people. Those in sub-Saharan Africa are largely

clustered in just nine of 46 countries, and they're mostly small, with operations that do not meet international standards. By comparison, China and India, each with roughly 1.4 billion persons have as many as 5,000 and 10,500 drug manufacturers, respectively. And the sub-Saharan market's value is still relatively small, at roughly $14 billion compared with roughly $120 billion overall in China and $19 billion in India (Should sub-Saharan Africa make its own drugs?," McKinsey & Company, January 2019).

Measuring Market competitiveness of the pharmaceutical industry
Measure of Pharmaceutical competitiveness can be done through either concentration ratio or Herfindahl Hirschman index (HHI). Concentration ratio is defined as the market sales or production accounted for by a certain number of the largest firms. Example the are the four firm ratio, eight firm ratio etc. A low concertation index implies that the market share of a particular industry is low meaning the firm is in an envelopment that is very competitive. A high concentration index implies that the firm is facing very low competition in it's in the market share.

HHI is a more in-depth measure of industrial competitiveness since its concentration index incorporates the size distribution of the firms. It is found by summing the squares of the market shares of each firm. The larger the HHI the lower the competitiveness of the firm in the industry. If the value under HHI approaches 0 it is a case of existence of very tiny firms corresponding to a near competitive market or perfect market. If on the hand the HHI approaches 10,000 it a case of Pure monopoly (10^2).The table xxx shows the competitiveness of pharmaceutical industry based on concentration and HHI index using US Bureau of the census 1992.

As it can be observed; using the two indicators pharmaceutical firms appears to be more competitive than many industries. The concentration ratio shows the percentage of the first four and the first eight leading firm combined to the total market. In the case of Pharmaceuticals the % of both C4 and C8 is quite low compared to all the other firms. This suggests that the industry is very competitiveness in the industry as per the

150

two classifications. Again, the HHI also low compared to majority of the other selected industry.

It is important to note that in case where pharmaceutical firms drugs has no close substitute, the HHI is expected to be close to 10,000 with concentration ratio of close to 100. Schweitzer (1996) illustrated a case of drugs used to control hypertension in four largest firms as having a concertation ratio of 91 implying that firm was nearly very close to pure monopoly see table 4.2 for clarity

Table 4.2 concentration ratio and HHI

Markets for Pharmaceuticals...

• Example of concentration ratio and HHI

Industry	Concentration ratioC4	Concentration ratioC8	HHI	N
Pharmaceutical	26	42	341	583
Cereal breakfast foods	85	98	2,253	42
Paints	29	43	305	1,129
Aircrafts and parts	79	93	2717	151
Passenger cars	84	91	2676	398

Barriers to entry in pharmaceutical industry

A barrier to entry is any factor that impedes the entry of new firms into an industry or product market. Pharmaceuticals firms uses different method to ensure that the market for their products is closely monitored and restricted by the use of the following methods;

Use of patent protection provided by government. A patent forms a legal barrier. This constitute an economic barrier of entry to the pharmaceutical industry. They ensure they have many variations of the product that remains patented.

Advertising and promotion: unlike the usual advisement for typical consume goods, Pharmaceutical firms promote their products through physicians rather than patients, by using 'detailers' who directly visits physicians. This technique helps at increasing brand royalty. Detailing is associated with questionable financial arrangements that encourage the physicians to prescribe a particular product. They also use other forms of advertising is through print media, bill boards, TV, radio, journal of medicines etc. the impact of this strategy as an economic barrier is not yet known though there are perception that these method are prone to misinformation.

Registration and Regulation of Drugs: Approval process for new drugs is costly and time consuming. A new firm will find it very expensive to have its product cleared for the market. This approval process requires a lot of financial and expert resources. For example, the statistics show that only 1 out of 5000 to 10,000 chemical compound screened is approved as drug, and of those reaching the market stage only 3 out 10 ever become profitable. Hence this is key barrier.

How are pharmaceutical industries regulated in your country?
This is one of the industries that is heavily regulated of all the other industries. Government regulate most firms for worker safety and health concerns, but pharmaceutical products are subject to further oversight beyond the measurements and brandings. The major reason being that drugs can have life threatening effect on the whole population. A live case is the USA, thalidomide drug that was used as tranquilizer widely in Europe, it was found to cause a severe defect in babies who born with deformed limbs.

Out of these experiences most countries have set organizations that specifically target pharmaceutical industries (mention such a body in your country). The bodies have come up with legal framework governing the production, sourcing, distribution, prescription, exposure, quality, testing etc. The legal framework is guided by some defined act, which in most cases delays provision of new drugs to the general population. For example, development of new drugs could take over ten years. The drug has to undergo lengthy and complex process aimed at testing safety, efficacy and effectiveness. It has to undergo a minimum of three phases i.e.

- ✓ Phase 1 – testing within a small group to ascertain safety and dosage
- ✓ Phase 2- involve large number of subject to ascertain efficacy
- ✓ Phase 3- conducted on thousands of patients to ascertain safety, efficacy and effectiveness.

Most countries in the world have drug and poison board to oversee that the framework on drugs registration and ethical processes on drugs formulation are followed. In USA they have Food and Drug and Cosmetic Act 1938.

Price setting under pharmaceutical industry

Pharmaceutical companies are painted as exploiting patients through patents and other strategies that reduce competition. Statistics from the books of accounts of these firms have found that they enjoy very high profits compared to other firms. Due to their barriers of entry, they have enjoyed an imperfect market that guarantee them to price their products following the in the spirit of imperfect market structure. This has supported downward demand curve which is fairy elastic though departing from the monopoly pricing.

The reason why pharmaceutical demand curve is not as elastic as in the case of a monopoly can be explained by the following factors:

1. Though patients with fixed copayments are not sensitive to higher prices and thus the expectation is that they should face a somehow perfectly inelastic demand curve, they could be affected by comprehensiveness of their health insurance

coverage. Maybe the cover could be having limited coverage on drugs or not at all. Hence, they are likely to substitute generic or over-the –counter product as drugs price increases.

2. Due to high prices some may reduce utilization by not complying with the medication. They will buy less due price change hence affecting price elasticity of demand of the pharmaceutical products.

3. The pricing inelasticity of demand curve can also be affected by purchasing decisions of the managed care organization and other large buyers, such as hospitals which could be sensitive to higher prices.

Market segregation or price discrimination under pharmaceutical industry

One of the main features of pharmaceutical firm is its ability to play a third –degree discrimination or market segmentation where different groups of buyers are charged different prices. Pharmaceutical firm can adopt price discrimination if it has the potential of segregating the markets. Under market separation the firm can charge different prices and thus be in a position to maximize profit. Price discrimination will depend on the characteristic of the market. Market characteristic will determine the price elasticity of demand. If one market has higher income and better insurance cover, the demand will be is relatively inelastic holding compared to a market with low income and inadequate health insurance. Figure 13 how price discrimination can occur given two hypothetical markets. There two markets M1 and 2

Market 1 has a more inelastic demand curve indicating that the firm can sell the drug at much higher price. On the other hand, market 2 has a fairly elastic demand curve indicating that the market is very sensitive to price change and thus a lower price will be required in order to increase sale. The conclusion is that even if the marginal cost will be the same in both markets the price will be higher in the market with less-elastic demand. Pharmaceutical firms will study the character of the market and formulate prices that will ensure that demand is maintained at a level consistent with market characteristic. For example, charging lower prices to hospital and managed care than

retail pharmacies due economies of selling and sensitivity of substitute that the hospital and managed care get.

Figure 4.13 Price discrimination

Price discrimation

Assignment

a) Take the case of a health market for pharmaceutical products and one for health workforce and consider to what extent they meet conditions for market optimality characterized by perfect competition.

b) What are the different reasons for observing variations in Physicians fee? Which reason are compatible with a completive model? Which reason are compatible with a monopoly.

c) Evaluate each of the advantages and disadvantages to physicians for joining a group practice as compared to in solo practice.

d) Explain the role of Government in the health care market. Why should government intervene in the health care market

Question for Discussion:

Consider the regulation of the health system in your country.

(i) Which bodies acts as a regulators and for what purpose?

156

(ii)   What means do they use?

(iii)   Whose interest do they generally represent?

(iv)   How effective are they in your assessment?

(v)   What alternative regulatory strategies might work better (e.g., positive incentives, or partnership arrangements)

References

Nicolson, W. (1985). *Microeconomic Theory, Basic Principles and extensions.* 3rd ed. Drden Press Holt, Rinerhart and Winston. ISBN 4-8337-0244-4

Varian, H.(1992) Microeconomic Analysis. 3rd Ed. W. W. Norton and company Ltd 10 coptic street, London. ISBN 0-393-95735-7

Zweifel,P., Breyer F., Kifmann, M.(2009). *Health Economics*, 2nd ed. DOI10.1007/978-3-540-68540-1_1, Springer-Verlag Berlin Heidelberg.

Folland, S., Goodman, A. and Stano, M. (2016). The *Economics of Health and Health Care*. Pearson International Edition Seventh Edition. Routledge. ISBN-10: 1292020512 or ISBN-13: 978-1292020518

Rice, T.H. and Unruh, L. (2015). *The Economics of Health Reconsidered*. Chicago: Health Administration Press 4th Edition

Feldstein, P.J. (2012) *Health Care Economics*, 7th Edition, Cengage Learning Publishers Inc. ISBN-10: 1111313261 | ISBN-13: 9781111313265

Witter, S., Ensor, T., Jowett, M. and Thompson, R. (20015. *Health Economics for Developing Countries: A Practical Guide*. KIT Publishers, 2015. ISBN-13: 978-9460221316 ISBN-10: 9460221319

Jack, W(199).   *Principle of health Economics for developing countries*. World Bank Development studies. ISBN 0-8213-4571-0
 Levit, Eugene M., and Douglas Coate (1982). The potential for using excise tax to reduce smoking. Journal of health economics I. 121-145.

Michael Conway, Tania Holt, Adam Sabow, and Irene Yuan Sun. (2019). Should sub-Saharan Africa make its own drugs?**,** McKinsey & Company

Taqdees Fatima, Shahab Alam Malik, Asma Shabbir, (2018). Hospital healthcare service quality, patient satisfaction and loyalty: An investigation in context of private healthcare systems", International Journal of Quality & Relia

Uwe Reinhardt, (1972). A Production Function for Physicians Services. *Review of Economics and statistics. P. 63*

Mahlathi, Percy and Jabu Dlamini. 2017. "From Brain Drain to Brain Gain: Movement of Medical Doctors in the South African Health System." Pretoria: African Institute of Health and Leadership Development. Accessed 10 August 2018. https://www.who.int/workforcealliance/brain-drain-brain-gain/17-304-south-africa-case-studies2017-09-26-justified.pdf?ua=1.

Obinna O., Jenny M., Benjamin U., Francisco M., Akin A., Thomas B., Roger W., Simon D. (2019). Identifying Key Challenges Facing Healthcare Systems In Africa And Potential Solutions. International Journal of General Medicine. Volume 12
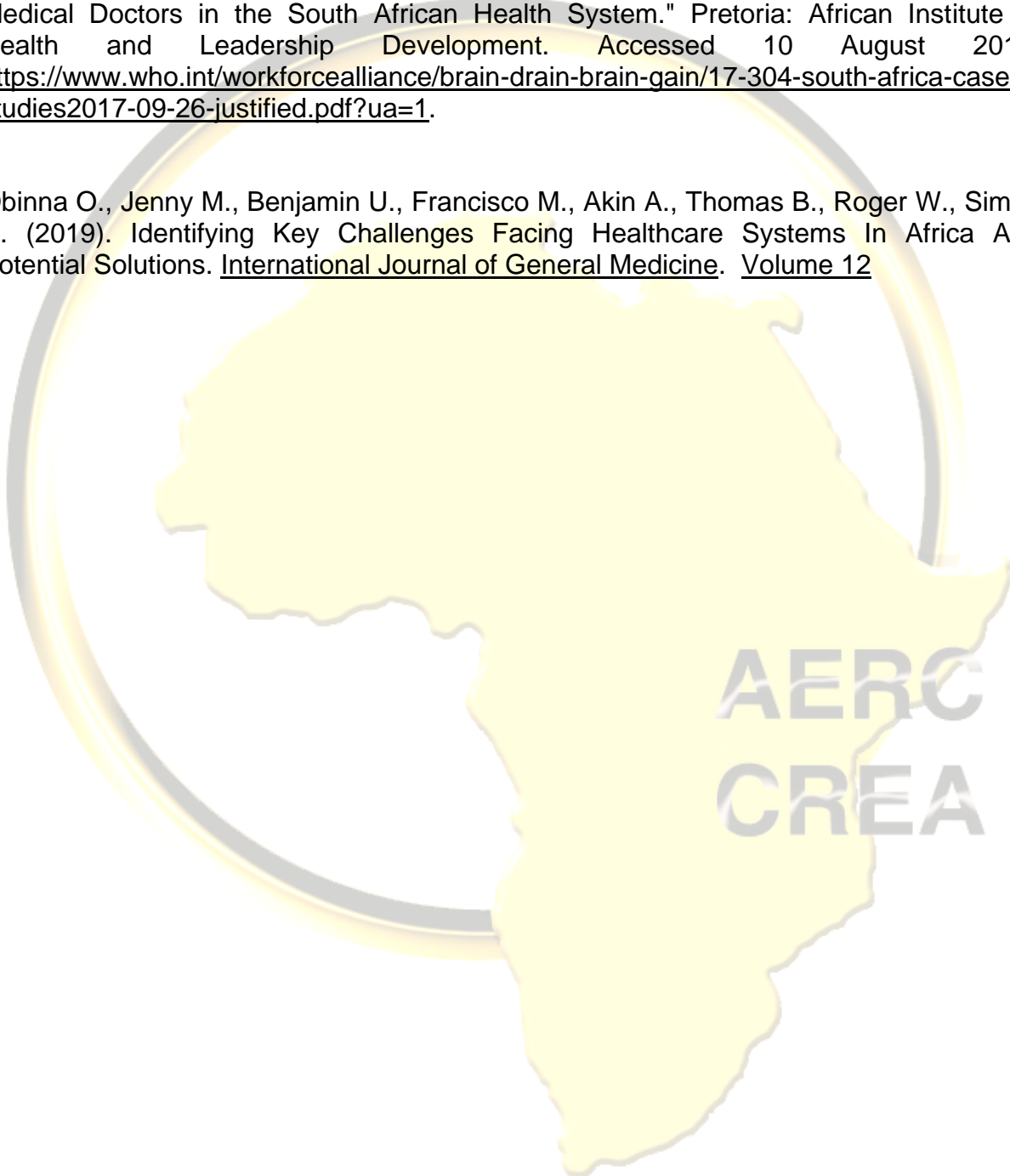
5   Health Insurance
Learning Objectives

By the end of this chapter you should be able to:

a)   Apply the model of health insurance in analyzing the health insurance markets.

b)   Develop the demand and supply models

c)   Evaluate the factors affecting demand and supply of health insurance

d)   Demonstrate the occurrence of market failures in health insurance market and propose solutions for the respective failures.

e)   Assess various available payment and reimbursement mechanism in healthcare market

Theory of Health insurance

For any discussion of health economy health insurance is critical. In developed economies health insurance is the cardinal method that people for their health care expenditure meaning consumer of health do not pay directly for their health. Rather insurance companies or other programs indirectly for much of the care with consumer paying directly a portion of the bill depending on the health insurance plan.

Why should consumer of health entrust somebody else to pay for their healthcare?

We will find this through the theoretical presentation of health care model. We start the model by arguing that Insurance exists due to uncertainty. Unhealthy day remains uncertain. Meaning illness does not strike with certainty. We do not know when we can fall ill as much as we could know our income or endowment. Hence, we need insurance to avoid uncertainty of ability to pay for our medical expenditure in future. Thus, the model proceeds as follows.

159

In health, individuals are not aware of their future health

It is assumed that we have two states

State s_1 –ill

State S_2 -not ill

Both states are mutually exclusive-they cannot occur at the same time from our model formulation we have:

P -give the probability value. Probability of being ill will determine whether individual buy insurance or not. To buy an insurance one has to pay a premium. The premium here is defined as a "actuarially fair" premium (considered for technical service purpose only).This is an assumption that the price of the premium is equals to the expected returns. This premium assumes that insurance companies are not incurring any expense like administration cost. The premium helps in the construction of the probability.

We can show that the expected wealth should be equal to

$$E(w) = p(w - m) + (1 - p)w$$

Where

W= initial wealth

M= the loss due to illness

P=probability

Hence expected wealth is a weighted average of the probability that one get sick which can be re-written as

$$E(w) = pw - pm + w - pw$$

Note that when pw cancels out you have W-PM as the expected wealth when sick. W-pw gives individual wealth if he did not take any insurance while the expected loss from sickness is equal to pm which can be represented as

$$pm = \pi$$

Where π Is equal to an Actuarially fair premium or policy. This mean that the expected value of wealth without insurance is the same as when one individual buys premium.

Hence a risk averse individual will prefer to pay a fair premium in order to avoid paying for future loss. Individual would prefer to even pay more, say like $\pi + \alpha$

α- here represents the risk premium which a risk averse person would be willing to add. Thus, it would be at the interest of the insurance company to charge a fair premium. If one assumes that there are n-individuals in a population and if all paid a fair premium and all make claims for compensation then we expect that that our Expected premium income=expected payments claims

In general insurance company will accept the fair premium plus an amount L or simply L is the loading factor which is included to pay for the transaction cost incurred by the insurance company.

For the insurance condition to operate we state that

$$L \leq \alpha$$

If this condition holds the insurance company can operate efficiently. If this condition holds the insurance company can operate efficiently. On the other hand individual can take insurance or self- insure if

$$L \geq \alpha$$

Existence of Peculiarities of health good suggest that some targeted welfare measures are always welfare-enhancing (e.g. taxation). But it does not always justify either the monopolization of supply through national health services or financing health care services wholesomely through tax Revenue. In essence, health goods should be paid by individual consumers. The problem with that assumption is that cost of illness occurs stochastically, meaning individuals have to hedge their income against inevitable uncertainty by buying health insurance.

Question: Can private insurance market deal with this demand optimally? Or do we need public intervention?

To answer this, we need to determine how individuals benefit from an insurance market by setting up an insurance model as the basis for this analysis

161

The Basic model of health insurance

Assumption of the model:

a)   An individual sick with probability---π

An individual remains in good health with probability 1- $\pi(0< \pi<1)$. When ill the individual require medical treatment to restore complete health. It assumed that the Cost of treatment takes on fixed value M. We already know that health insurance is characterized by premium P and insurance benefits I. It also assumed that individual earn exogenously given income and he derives utility from disposable income. This type of utility is a Bernoulli utility function or Von-Neumann Morgestern utility function meaning that utility function is defined on sure amount of money.

The basic model is then formulated as follows

We start by assuming that if the individual is healthy disposable income can be expressed as follows

    i) y_h= Y-P

Again, we show that incase of sickness disposable income is given as

    II) y_s=Y-P-M+I

Hence our model becomes:

$$EU = (1 - \pi)\mu(y_h) + \pi\mu(y_s)$$

$$y_h = Y - P = Y - \pi I$$

$$y_s = Y - P - M + I = Y - M + (1 - \pi)I$$

$$\mu'\left[Y - \pi I^*\right] = \mu'\left[Y - M + (1 - \pi I^*\right] \Rightarrow I^* = M$$

Where:

EU = Expected utility of the individual-

u(y) = Utility function---

y_h = Disposable income when healthy--

y_s = Disposable income when sick--

π = Probability of being sick--

Y - =Gross income--

M =Cost of treatment---

I = Insurance benefit

P= πI = actuarially fair health insurance

We note that individual maximizes expected utility which is given by

$$EU = (1-\pi)\mu(y_h) + \pi\mu(y_s) = (1-\pi)\mu(Y-P) + \pi\mu(Y-P-M+I)$$

Assuming that insurer incurs no administrative cost and that competition is perfect then insurer offers actuarially premium. Then P= πI *i.e.* premium is equal to the expected benefits

With P= πI , the optimal quantity of insurance benefits *I* can be found by solving

$$\max EU = (1-\pi)\mu(Y-\pi I) + \pi\mu(Y-M+(1-\pi)I)$$

The first order condition is given by

$$\frac{\partial EU}{\partial I} = (1-\pi)\mu'\left[Y-\pi I^*\right](-\pi) + \pi\mu'\left[(Y-M+(1-\pi)I^*\right](1-\pi) = 0$$

Re-arranging the first order condition gives

$$\mu'\left[Y-\pi I^*\right] = \mu'\left[(Y-M+(1-\pi)I^*\right]$$

This expression mean that the marginal utility must be the same in both states that when health and when ill. Therefore, full insurance with *I*=M is optimal and the individual has certain income shown by

$y_h = y_s = Y-\pi M$. This result characterizes an efficient market. Therefore, full insurance with *I*=M is optimal and the individual has certain income

$y_h = y_s = Y-\pi M$. This result characterizes an efficient market.

Individual's initial situation is given by point Y where $y_h = Y$ and $y_s = Y-M$.

This simply means that the purchase of insurance changes state-contingent income to

$y_h = Y-P = Y-\pi M$

$y_s = Y-P-M+I = Y-M+(1-\pi)I$

Given that M=I, you can prove that $y_h=y_s$

Demand for and supply of health insurance

Consumer demand for health insurance implies amount of health insurance he/she is willing to buy at different premium from health insurance market. From our model of health insurance discussed, the goal of maximizing total net benefit provides the framework for understanding health insurance choice. We note that a person benefit from health insurance only when ill though there is the unobservable satisfaction one get that is tied with reduced uncertainty. It is also notable that the marginal benefit when well will decline as additional insurance coverage is purchased as we apply the law of diminishing marginal utility. In contrast, the marginal costs when well rise as additional insurance is purchased. The solution to purchase is that the marginal cost equals the marginal benefit.

The key assumption as before is that the individual wishes to maximize his or her utility.

Since individual does not know whether they will have illness or not and how much they loss in case of illness, it means that they maximize expected utility rather than actual utility due to uncertainty. In this regards the individual can choose between two alternative courses of action.

a) Purchase insurance thus incurring a small loss in the form of insurance premium

b) self –insure, meaning facing the small possibility of a large loss in the event that illness occurs or the possibility that the medical loss will not occur

We can illustrate the willingness of the individual in the purchase of health insurance by use a hypothetical assumption (see Feldstein 4th edition 1993)

Using the Figure 5.1 we and assuming a case that;

a)   if illness occurs it will cost the consumer $8.000.

b)   Individual wealth is $10,000(W3 and U3),

c)   After illness wealth will be 2000(W1 and Utility U1**)**

If we assume that the probability that illness occurs is 0.025 implying that pure premium of insurance to cover $8,000 is loss is 0.025X$8,000= $200 (note pure premium or actuarial premium is a function of both size of expected loss and its probability of event occurring).Pure premium reduce wealth by a certainty of $200 if the individual pay premium. Wealth is now $ 9,800(W2 and U2). The choice facing individual as also shown in our Figure 5.1 therefore are:

(i)   Purchase insurance for 200 and move to a lower level of utility, U2

(ii)   Not to purchase insurance and have a 2.5 percentage chance that he/she will incur an $8,000 in case an illness occurring (utility at U1) or remain at the same level of wealth($10,000 with utility U3) if illness does not occur.


What is the relationship between actual utility and expected utility in the purchase of a health insurance cover? This is again shown in our two figures. Actual utility will follow the concavity assumption of the total utility curve of wealth which for a "rational" individual is known to decrease over time, meaning additional wealth has somewhat lower marginal utility as shown in our first figure. On the other hand expected utility will be determined by weighted probability and the trend will be on decreasing linear fashion (**shown by Feldstein page 112-115**).The expected marginal utility due to wealth will decrease with increase in probability of an illness occurring. Again note actual utility curve (decreasing marginal utility with respect to wealth) is always above the expected utility. What this mean is that individual will always buy insurance at its actuarially fair value(pure premium). If the expected utility for not purchasing insurance is lower than

165

that of purchasing insurance, as shown in Figure 5.1 the individual is likely to purchase insurance. The straight line represents expected utility for different probabilities that illness will occur. Note the lower the probability that illness will occur the closer the expected utility will be to the farthest to the right on utility curve. While high probability of illness occurring will be associated with expected utility further down to the left at U1 on the curve.

Figure 5.1 Relationship between Total Utility and Wealth

Relationship between total utility and wealth

Actual utility

Expected Utility and actual utility
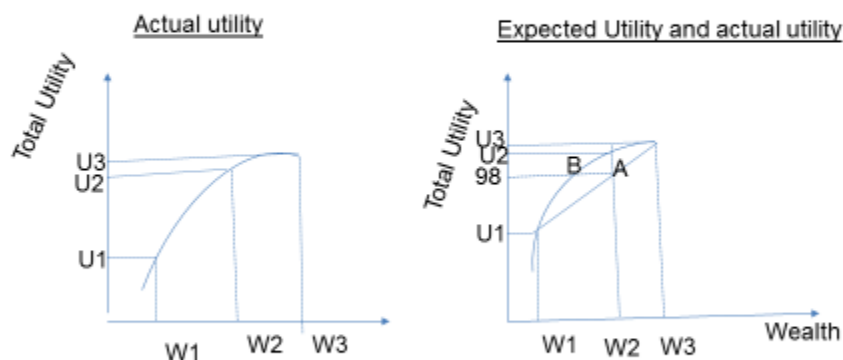
What message do we get from this illustration that can guide the understanding of demand for health insurance?

Answer! We note four things

1. Appropriate amount of health insurance coverage arise when the marginal benefit of health equals the marginal cost of purchasing that insurance.

2. The assumption is that 100 percentage coverage arises when the medical expenses is equated to the insurance premium net of any loading cost.

3. If insurance administrative cost is added on the actuarially premium, consumer would demand less than 100 percent.

4. Adding an administrative cost to the pure premium would cause the total price of the last units to be greater than their marginal benefits.

Highlight what would facilitate or not the willingness to purchase health insurance above the pure premium.

A case where illness or loss is becomes almost certain to occur, the person can save for the event instead paying the same amount (equal to an insurance pure premium) plus an addition amount to cover other insurance company cost. In case of near-certain events such as annual medical or dental checkups, the probability of purchasing health insurance policy that is above pure premium will be very low. One can easily plan for this cost without necessarily getting into expensive insurance premium.

The magnitude of the expected loss. If the expected loss from an illness is expected to be very high, the person is likely to increase the probability of purchasing insurance cover that is inclusive of pure premium and insurance company cost. On the other hand if loss is perceived to be small the demand for health insurance that include pure premium and company cost will be low

Concluding remark here:

A person is more likely to insure against events that have a greater magnitude of loss than against events with smaller possible losses. Similarly, the amount above the pure premium that person is willing to pay is less for small losses than for large losses given the same probability. **What is the implication of this assumption to today world?**

With the cost of medical skyrocketing and the probability of loss increasing this has caused an increase (shift) in the demand curve for insurance

Let us now turn to look at factors known to affect the demand for an insurance policy.

Characteristic of the individual in terms of risk averseness. If individual is faced with a positive marginal utility curve that is decreasing he or she would not to remain at the same level of utility. This like saying as income increases for the individual increases he or she exhibits an increasing utility with a decreasing marginal utility. This individual would not want to loss utility in future and such he or she could be willing to pay pure premium with administrative cost because of his or her risk averse characteristic.

The probability of the event occurring: for those events that have very low and very high probability of occurring, individuals are less likely to pay above the pure premium compared to events with intermediate probability of occurring. One would ask but why is it that one will not be willing purchase health insurance when the probability of occurring is very high? The answer is that the expected premium will be too high to an extent that their marginal cost will be much higher than their marginal benefit.

The magnitude of the loss: The larger the magnitude of the loss, the greater will be the amount above the pure premium that the individual is willing to pay.

The price of insurance: The higher the price of insurance due to loading cost (the amount above the premium cost), the fewer will be the events the individuals will insure against.

The income of the individual: The size of a person's income and wealth affects the amount above the pure premium they are willing to pay for health insurance. At both low and high incomes the marginal utility of income is either high or low, so that such persons might want to self-insure.

Tax treatment: High income has another effect on the demand for health insurance in that insurance policy has a fringe benefit for it is not considered for tax, hence a motivation to acquire an insurance policy. The tax subsidy for health insurance lowers the price of insurance to those with high income. Notable is as incomes increase and people moving into high income brackets, there is greater incentives for them to demand fringe benefits from their employer rather than increases in their cash incomes.

Occupation and groups. The occupation where a person is matter a lot in terms of demand for health insurance. This spell on out whether an individual is employed as part of a large employee group which means the group can purchase a group health insurance. Group health insurance policies are sold at substantially low prices. This is due to the fact that administrative costs are deemed low. Again under group members there is less likelihood of adverse selection. Unlike individuals' insurance policy where individuals plan to use cover in near future, members of a group join because of being in a job rather than due to because they were ill and wanted insurance.

Change and advancement in health technology: Progress in technology has made it possible to treat diseases that were initially not treatable. These treatments are costly. Given that these treatments are available there is increased probability that individuals might require them. The magnitude of loss is increased as well as the probability of loss occurring. This is likely to cause an increase in the demand for health insurance.

Discuss technology in treatment as a derived demand of hospitals infrastructure

Question: Is there any Welfare implication on the theory of demand for health insurance? Motivate your answer
Answer: Yes!

We can explain this by upholding that, If the marginal benefit is less than price of the insurance policy, consumers will be worse off because they equate marginal benefits of good and service to the price (which reflects the marginal cost of producing those goods). They will cut the consumption of those goods and services whose marginal benefits are less than price. They increase their utility by using the fund saved to purchase other goods and services whose marginal benefits are higher than price. If consumers are forced to purchase a good whose price is greater than its marginal benefit, they end up being worse off than before. This is the case with comprehensive insurance coverage where individuals are expected to insure for all their medical expenses.

Hence the argument is as long as different forms of coverage are divisible and do not have to be sold together, consumers would be better off with some coverage or none at all. The welfare implication of mandatory insurance coverage that cover all medical losses is that some consumers will be worse off than if they had a choice self-insure in those situations

Supply of Insurance

To know how much insurance policy will be supplied in the market we start by looking at how premium are set in a competitive market. We are assuming that insurance companies are operating under a perfectly competitive environment where abnormal

profits are not observed since once a company makes abnormal profit other competing firms will come to competes for that profit. The market then can only operate under a normal profit. Let use a model of insurance market to demonstrate the amount that insurance company are willing to supply.

Competition and Normal profits

We start with a hypothetical example;

Assume Jane is buying insurance in a competitive market and she faces a potential illness with a probability of 0.05. She sought to buy an insurance cover of $500 and her insurance company was charging her 10% premium which is $500*0.1= $50. Assume that it cost insurance $5 dollar to process each insurance application. Her insurance company profit can be defined as

Note cost of illness to the insurance company= $5+$500

Profit= Total Revenue- Total Cost.

Profit= $50 –(Probability of illness X cost if ill)-(probability of no illness X cost if no illness)

Profits = $50- (0.05 X 505)-(0.95 X $5)

Profits= $50-$25.25-$4.75

Profits (premium=10%) = $50(revenues)-$30(Costs)

These are positive profits and remember in a perfect competitive markets no firm should make positive profits. Other firms will come to compete for this profit.

Lets find what happens when lower the premium to 8% assuming Jane assumption are held constant and that firm cost structure is unchanged.

We note that the cost side of the equation will remain the same hence our profit will be

Profits (with 8% premium) =$ 40(revenues)-$30=$10

The profits are still positive.

It easy to show that at 6% of premium the normal profit will be attained. For the insurance we can put that the revenue per policy is, aq, where a is the premium, infraction term. The cost of policy in terms of payout is the probability of payout,

p, multiplied by the amount of payout, q, plus a processing cost t, which is unrelated to the size of the policy(assuming it cost no more to administer $10,000 policy than a $500). So:

Profit=Revenue-Payout=$aq - (pq + t)= aq-pq-t$
With perfect competition, profits =0,
So:
$0=aq-pq-t$
Solving for the competitive market premium a, as ;
$a=p+(t/q)$

the expression shows that the competitive value of a equals the probability of illness, p, plus the processing (or loading) costs a percentage of policy value, q, or t/q. if load costs are 10% of policy value , q, then (t/q) = 0.10. Hence, in equilibrium a=(0.05 +0.10). Hence the premium for each dollar of insurance, q, of is $0.15. if insurer charge less, they will not have enough money to pay claims. If they charge more, they will have excess profits that will attract other firms to come and compete for this profit in the market.

Note that this is the basic on how insurance determines supply through premium setting. There several ways this can be improved by adding other assumptions as seen Zweifel et al., Health Economics (2009) second edition. Page 163-195

An empirical case of Health insurance
Let us now turn to RAND Health Insurance Experiment (RHE) on the demand for health care
RAND Health Insurance Experiment (RHE)
RAND Health Insurance Experiment (RHIE) data form RAND Corporation in USA has been used to analyze demand for medical services in the face of co-insurance policies. RAND data comprised of one of the largest randomly controlled economic experiment.

The data was designed to test the effect of alternative health insurance policies on the demand for health care and on health status of both treated and non-treated groups. In the experiment, participants were randomly assigned to one of the insurance plans for a period of five years. The insurance plans included:

a.) Pay no out-of-pocket for their health insurance bills- what is called free plan

b.) Paying 25% of the health insurance bill

c.) Paying 50% of the health insurance bill

d.) Paying 95% of their health insurance bill

The key findings from this experiment were:

- The greater the portion of their health care bill that individuals are required to pay, the less health care they choose demand

- The fully insured purchased 40% more than those who had to pay 95% of their health care insurance bill

- The 40% increase in health care utilization had little or no effect on health status for the average adult according to Newhouse et al. (1993)

- Children under the cost sharing scheme consumed up to 1/3 less health care. But the reduction in care was not significantly related to health status (Valdez et.al 1985)

- As coinsurance rose, overall use and expenditure fell, for adults and children combined

- Compared to free care plan, a coinsurance rate of 25% resulted in a 19% decline in expenditures

- While higher coinsurance rates of 50% and 95%, resulted in over 30% decline in expenditure. The probability of poor children in cost-sharing plan having an episode of treatment was 44% less than children in the free plan

- The probability of non-poor children having episode of treatment was 15% less than children in free plan

- Poor adults in the cost sharing plan had a 41% lower probability of seeking treatment than adults in free plan, while for non-poor adults it was 29% lower

Main conclusion from RAND experiment:

This suggest price and income is a key determinant in the demand for health care. The finding also suggests that poor individuals are at a greater risk of not getting treatment under cost sharing plan. The implication is that government proposal on National health insurance policy should consider cost sharing bearing in mind the level of individual income. The findings should build on findings coming from other countries in order to strengthen their policy argument

Market Failure in Health Insurance:

Adverse selection and Moral Hazard

Insurance theory assumes that Insurance companies are going to operate efficiently through normal profit margin. Relation between wealth and risk aversion will depend on loading factor coverage and purchase. This again will depend on an event that has a high probability of occurring. Events are classified into two: high probability event (common cold) and low probability event like surgery. People left on their own would purchase insurance on the basis of their occurrence. In this case the operation of a free market would lead to distortion or market failure. There are two main market distortions that confront insurance company, that is, adverse selection and Moral hazard.

Let us discuss each in turn;

174

Adverse selection:

Adverse selection occurs in situations where individuals are in position to assess the risk they face in a situation better than an insurance provider can. In this case individuals who know they have small risk will not buy insurance coverage-it is too expensive given what they know-where high risk individual will purchase coverage. This situation occurs when individuals are able to purchase insurance at a rate below actuarially fair rate plus loading factor. It arises due to asymmetric information by insurance companies.

We use a theoretical approach to illustration how adverse selection arises:
We start by assuming a population of two sub-groups one with high probability of illness and the other sub-group with low probability of illness
N=H+L
H-High risk (children, pregnant mothers, aged chronically sick and
L- low risk (age24-50) are defined in term of frequency of illness associated with each. If insurance was to operate fairly the company would charge a fair premium according to risk associated to each individual. Assume again a high-risk premium is π_h and low risk π_L. Here we also assume that the insurance company does not know which individual is L or H, this is a case of asymmetric information. Under this assumption on would expect that the insurance company would charge a uniform premium or what is referred to as *community rated or based premium (CRP)* which is arrived from the average of the two population assumptions, that is:

π_c is CRP which is an average premium of the two sub-population

Will the company be able to operate if it were to charge CRP?
The answer will be Yes, if the company is able to cover its operational cost. This is possible if the individuals are willing to expose themselves by providing correct population sub-group. With correct sub-population captured the company will use the formula and get the right premium that will make it operate efficiently.

175

On the negative that No response the company will not be able to operate with CRP: the reason is due to the fact that after the average is computed the scenario would be $\pi_c > \pi_L$ thus the low risk would find premium high and not purchase while the high risk demand would very high because $\pi_c < \pi_H$ thus combining the two we discover that insurance company would only cover those in the high risk areas. If copy is covering only high risk premiums it means that the transaction cost will be more than the receipts thus making company very hard to operate.

For insurance to operate then it would be crucial to make insurance compulsory.
But even this is likely to be faced by challenges because we still have H and L risk in the population. In a free market, insurance company will go for the L and thus avoid H. This is a case of adverse selection if the individuals are asked to review themselves. Individual know that if they disclose themselves as high they are not likely to be covered or they will be required to pay very high premium.

If on the hand the company was a monopoly it would charge premium
$\pi_c > \pi_L$ this implies it would charge community rated premium that is higher than the one charged on the L probability sub-group.
If insurance copy was not monopolized experience rating premium would be better than CRP. In experience rating premium, people normally pay a premium directly related to their risk. This is only possible if insurance company are able to separate H and L

Is there any problem with experience rating premium?
The answer is to the affirmative or yes. One, if the H-group population constitutes the low income group, we expect that they will not be able raise or pay such high premium. Two, if the sub-group of high constitute the chronically ill, it is expected that the fair premium will be very high hence lowering the probability purchase by the chronically ill due to premium. The result of this formulation indicate that due to sample selection basis those with the highest needs are not likely to benefit from insurance policy.

Meaning the H risk population is likely to miss out insurance coverage creating even a bigger market failure.

So what can we do?

In general society care for such group on the premises that health care has some adverse externalities to the rest of the community. Thus, the best solution is to incorporate the market solution plus some subsidy or government fiscal policy. This can be done in two ways.

a)   Selective subsidy in terms of cash transfers to purchase insurance. The only problem with this kind of arrangement is the issue of fungibility which mean the recipient allocating these cash transfer to other urgent needs.

b)   In-Kind Transfer; this where the government buy insurance policy for the very poor like the elderly or those who are vulnerable. This has proved to be much better than cash transfer

c)   Another way to solve market distortion in the insurance market for health is to forget about use of market and follow either Bismark model or Beveridge model. According to Bismark model insurance is provided publicly. In this case individuals and employers pay premiums. This premium plus state subsidy are used to finance health care services. The premium is based on community rating. The insurance coverage is again compulsory and monopolized. On the hand Beveridge model health care finance is financed from general taxes hence one might want say it is not a genuine insurance scheme. It is compulsory and there is a monopoly suppliers. Example is UK where all the health services are being provided by National services.

Moral Hazard

Moral hazard problem arises when being covered by insurance against some event actually increases the expected value of loss. An insurance term that represents the

disincentives created by insurance for individuals to take measures that would reduce the amount of care demanded. It expresses the additional care demanded resulting from a decrease in net price of care attributable to insurance. In the study of medical insurance, it is well known that coverage of certain types of medical condition such as psychiatric or dental care increases medical care substantially. Once one is covered by health insurance the consumer medical expenditure at that point can be zero or very small depending on the health policy plan, that is demand tends to increase though it highly unlikely that it reflects the actual need. This situation causes market failure or distortions in the insurance health market. Unlike in adverse selection where probability is self-determined, in moral hazard individual can influence the probability of being ill and or not. In real case buying insurance lowers the service price. The extent of moral hazard will depend on the elasticity of demand for health care. If the demand is inelastic moral hazard will be rare (**EXPLAIN WHY**). In an example where we have frequent demand like demand for insulin for diabetes condition, case for moral hazard is rare because the health care has become a need than a want. If on the other hand demand is elastic moral hazard is highly likely. A good example is the demand for cesarean operations when the patient demand is due to "thy neighbor" policy principle (when cesarean is not done out of emergency but rather out of choice).

Figure 5.2 if demand for medical care were inelastic with respect to price, individual's demand curve in the event of illness would look like D1. In the presence of insurance the individual will demand according to D2 since he /she is operating on an elastic demand curve. The presence of some elasticities in the demand curve indicates that the individual's demand different quantities of medical care depending upon how much must be paid for the care. Since the insurance lowers the price of medical care to individuals, they will probably consume more care than if they had to pay the entire price themselves. It is this behavior of individuals that is called moral hazard.

Figure 5.2: The demand for medical care under conditions of moral Hazard
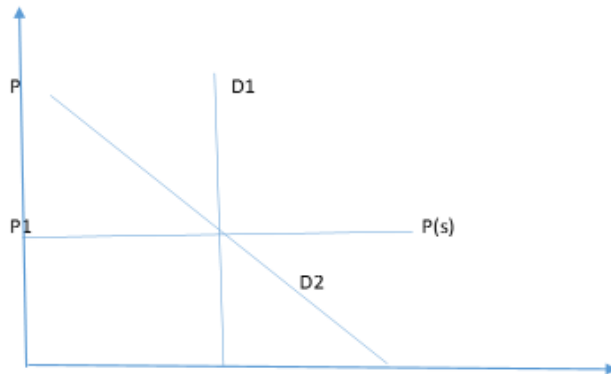
Given the fact that moral hazard is the insurance tendency to alter individual motivation to prevent a loss, then it implies that overconsumption of unmerited health care goods are consumed. This in turn affects expenses for insurance companies and consequently the cost of covering such individuals.

How can we control moral hazard in health care market?

Two key method has been identified

a) As result insurance company have introduced co-insurance as away of discouraging over consumption of health good. Consumers are expected to pay a certain proportion their medical cost to the health facility or the provider. This is what they call cost-sharing of a kind. In this case individual would pay like 20% and company pay the rest 80%. This tends to lower moral hazard.

b) Another that moral hazard can be controlled is by educating the general public about the danger of over-medicalization. Over-medicalization is known to jeopardize the person's body immune system.Individuals.one source of over medicalization is having a health insurance where individual demand more health care than they need due to perception of taking advantage of low prices.

Payment and reimbursement mechanisms

There are four basic model that are used in the provision of health care to the general population. These basic models are aimed at, keeping people healthy, treating the sick, protecting families against financial ruin from medical bills and addressing adverse selection problem. The four basic models include:-

- The Beveridge Model

- The Bismarck Model

- The National Health Insurance Model

- The Out-of-Pocket Model

❑ The Beveridge Model

This payment model was named after William Beveridge the daring social reformer who designed Britain's National Health Service. According to this model health care is provided and financed by the government through tax payments, just like the police force or the public library. Under this model the suitability is guided the fact that many, but not all, hospitals and clinics are owned by the government. Another reinforcing issue is that some doctors are government employees meaning they are payed through salary though there are also private doctors who collect their fees from the government. These systems tend to have low costs per capita, because the government, as the sole payer, controls what doctors can do and what they can charge. Countries using the Beveridge plan or variations on it include Great Britain, Spain, most of Scandinavia, New Zealand and Hong Kong

The Bismarck Model

This reimbursement model was named after the Prussian Chancellor Otto von Bismarck, who invented the welfare state as part of the unification of Germany in the 19th century. This model uses an insurance system called "sickness funds" This type of payment is usually financed jointly by employers and employees

180

through payroll deduction. Bismarck- health insurance mode of payment system is geared towards to covering everybody. This model is aimed at making ensuring that everybody is covered under its scheme without the motive of making any profits. In this model doctors and hospitals tend to be private which act as good incentive to making business outside salary arrangement. This is a good provider payment mechanism because the doctors and hospitals will work toward maximizing their profits as well as maximize the welfare of the citizens. This model is found in Germany, France, Belgium, the Netherlands, Japan, Switzerland, and, to a degree, in Latin America.

The National Health Insurance Model

This is a very popular payment system has elements of both Beveridge and Bismarck. It uses private-sector providers, but payment comes from a government-run insurance program that every citizen pays into. Since there's no need for marketing, there is no financial motive to deny claims and no profit. These universal insurance programs tend to be cheaper and much simpler administratively than American-style for-profit insurance. The single payer tends to have considerable market power to negotiate for lower prices. National Health Insurance plans also control costs by limiting the medical services they will pay for, or by making patients wait to be treated. Canada's system, for example, has negotiated such low prices from pharmaceutical companies. The classic NHI system is found in Canada, but some newly industrialized countries like Taiwan and South Korea, for example, have also adopted the NHI model

The Out-of-Pocket Model

This is very popular reimbursement mechanism especially in countries where insurance markets and more the health insurance markets are not very developed. It is also very common in countries where informal sector is very big. In this mode, you pay from your pocket at the time of treatment or when you are sick enough to be admitted to the emergency ward at the public hospital of. Developed and industrialized countries have developed health care systems these constitute like 40 countries out of 200 countries of

the world. The rest of the countries tend rely heavily on the out of pocket mode of payment. Most of the nations on the planet are too poor and too disorganized to provide any kind of mass medical care outlined earlier that is out -of –pocket system is the ultimate mode of reimbursing health providers. The basic rule in such countries is that the rich get medical care; the poor stay sick or die. In rural regions of Africa, India, China and South America, hundreds of millions of people go their whole lives without ever seeing a doctor. They may have access, though, to a village healer using home-brewed remedies that may or not be effective against disease. In the poor world, patients can sometimes scratch together enough money to pay a doctor bill.

Managed Care- payment and Reimbursement

What is managed care as a form of reimbursement of health care? And what is the motivation for a managed care and in what form does it exist?

❑ Managed care:

Managed care is any payment or delivery arrangement used by a health plan or provider to control or to coordinate use of health to contain health expenditures, improve quality or both. It is normally provided through a specific network of hospitals, Physicians and other health providers. There is considerable centralized oversight of resource use referred to as utilization review which occurs within the network. The plan tends to have a defined system of selected providers that contract with them. To large extent they are type of health insurance. They have contracts with health care providers and medical facilities to provide care for members at reduced costs. These providers make up the plan's network

The motivation of managed care payment hinges on its intention to reduce unnecessary health care costs through a variety of mechanisms, including:

a)   Economic incentives for physicians and patients to select less costly forms of care

b)   Programs for reviewing the medical necessity of specific services;

c)   Increased beneficiary cost sharing

d)   Controls on inpatient admissions and lengths of stay

e)   Establishment of cost-sharing incentives for outpatient surgery

f)   Selective contracting with health care providers

g)   Intensive management of high-cost health care cases.

We have several types and feature of managed care, with main one being, Employer-sponsored managed care which has Health Maintenance Organization (HMO), Preferred Provider Organization (PPO), Point Of Service (POS). We also have Medicaid managed care, managed care contracts with Physicians and managed care contracts with Hospital.

Let us look at some of the feature of the highlighted types of managed care.

Health maintenance organization (HMO): The HMO concept was promoted by the Nixon administration as a fix to rising health care costs and set in law as the Health Maintenance Organization Act of 1973. HMO would, in exchange for a subscriber fee (premium), allow members access to a panel of employed physicians or a network of doctors and facilities including hospitals. MO is a coordinated delivery system that combines both the financing and the delivery of health care for enrollees. In the design of the plan, each member is assigned a "gatekeeper", a primary care physician (PCP) responsible for the overall care of members assigned. Specialty services require a specific referral from the PCP to the specialist. On-emergency hospital admissions also require specific pre-authorization by the PCP. Services are not covered if performed by a provider not an employee of or specifically approved by the HMO unless it defines the situation to be an emergency. HMO that directly employs their own physicians are called staff model plans. In self-model plans physicians paid salary by the HMO.HMO plans that contract physician from a geographical spread out, independent solo, or small

group practice are called independent practice associations (IPA). Self-model plans and IPA assign Primary Care Physician as the gatekeeper.

Preferred provider organization (PPO): PPO is a managed care organization of medical doctors, hospitals, and other health care providers who have agreed with an insurer or a third-party administrator to provide health care at reduced rates to the insurer's or administrator's clients. A membership allows a substantial discount below the regularly charged rates of the designated professionals partnered with the organization. Preferred provider organizations themselves earn money by charging an access fee to the insurance company for the use of their network. PPOs can also contract with one another to strengthen their position in certain geographic areas without forming new relationships directly with providers. PPOs have no gate keeper, rather patient must pay more out of pocket if they choose to go outside the network. Hence, PPOs create financial incentives for subscribers to use network providers rather than outside the network for care. They contract with the physicians and hospitals generally to address the prices providers will charge the PPO. If Provider charges less than the average price in the market, he/she become the preference of PPOs contract. Providers always agree to undergo utilization review under the contract. It is notable that most PPOs require preadmission certificate for hospital stay and concurrent utilization review for such stays. Majority of them require a mandatory second opinion for a recommendation of surgery.

Point of service (POS) POS is a type of managed care health insurance plan. It combines characteristics of the health maintenance organization (HMO) and the preferred provider organization (PPO). It charges lower medical costs in exchange for more limited choice.in this managed care, coverage is greater when member use network providers. Under POS coverage is less generous when member use non-network providers. They assign each member a physician gatekeeper who must authorize in-network care before care is covered. In this managed care there is no

requirement for authorization to use a out-of-network services, but that care is covered on a less generous terms.

Medicaid Managed Care Plans (MMCP); these are targeted health insurance programs administered by the state government for qualifying low –income beneficiaries. These initiatives are focused on improving care for populations with chronic and complex conditions, families with dependent children and pregnant women who meet low-income criteria. It can be managed through HMOs that exist in area or the state can form their loosely structured provider networks which contract with selected providers for discounted services. MMCP require that in both case physician gatekeeper is a must for the control of utilization.

Managed care contract with the Physicians; Most HMOs and POS plans pay their network physicians on capitation basis. This fixed amount is actuarially determined. Physicians may be responsible for the costs of referrals laboratory tests, and hospital services. Hence the physician is exposed to a risk of either not making profit incase these additional turns to be more than the capitation per patient. PPO contract with physicians is mainly based on some discounted fees for various services that plan will pay physician being in the network. They usually get pre-negotiated fee as payment for the subscriber who has received service. Utilization review procedure are commonly covered under HMO, PPO or POs. Most managed care contracts require that a degree of record keeping for the client by the physicians. Some physicians limit the risk associated with capitation by obtaining stop-loss provision in their contract- Stop loss provision limit the physician's liability per enrollee by an amount of money say $25000 per year for a single enrollee. They can also purchase reinsurance against large loss.

Managed care Contracts with hospitals; HMO and PPO plans contracts with limited subset of the providers-physicians and hospitals) in the areas they serve. This arrangement has promoted price competition among hospitals that might otherwise lose

plan business. Some HMO have owned their own hospitals that their members use. Contract between individual managed care and hospital depends on:-

✓ Plan Characteristics, that is, is it a HMO, PPO, plan size, whether the plan serve several localities and how old the plan is.

✓ Hospital characteristics:- size ownership(whether nonprofit or for profit), location(suburb Vs city), teaching status, cost structure

✓ Market characteristics:- penetration and rate of growth of managed care plan

Empirical work: Zwanziger and Meirowitz (1998) examined the determinants of managed care plan contracts with hospital and found that:-

a) Managed care plan prefers contracting with nonprofit hospital, and preferring even public hospitals to for-profit ones

b) Contract with large hospital compared with medium and small hospitals, and contract more with medium hospital compared with small hospital

c) Hospital cost did not matter or was not significant, it did not affect contracts. This was surprising (could it imply that managed plan use prices as proxy for quality, or could be that public correlate cost low cost with perceived low quality)

Value-Based Purchasing (VBP)

What is Value-Based Purchasing (VBP)?

In healthcare, VBP links healthcare payments to improvements in the patient's individual care. It has an objective of reducing the amount of care given for the profit motive. It also rewards for the best-performing healthcare providers. Value-based

purchasing help to improve the quality and to some extent reduce the number of lawsuits due to poor quality and costs. It is intended to motivate care providers to perform their best on the job. It also helps create awareness especially when it comes to prevention of illnesses at the community level through Home Health- Value -based set up.

Home health value-based purchasing also keeps home healthcare accountable. We can cite at least two types of VPB as;

a) Hospital Value-Based Purchasing

b) Home Health Value-Based Purchasing

Hospital Value-Based Purchasing (HVBP)

Hospital value-based Purchasing works is aimed at rewarding the quality of clinical care provided to patients in the hospital setting. The value is based the treatment of the patient during the extent of their stay in the facility including how closely good practices were adhered to. It addresses the competitive edge of the facilities in terms of whether they perform better than the peers as well gauging how much it has improved since the start of the data collection period. The baseline data are collected before the patient's treatment commences. HVBP measures on everything from the conscientiousness of safety to the level of customer or patient-centeredness that the care was done.

Home Health Value-Based Purchasing (HHVBP)

These types of value purchasing are mainly found in developed countries where health provider traces patient or on a follow up bases to ensure they are following treatment protocol at homes. They not very common in developing countries though they are currently creeping through community health worker. In developing countries they are mainly linked with follow on HIV and TB patients to ensure that they are adhering to treatment procedures given from the health facilities. In developed countries they are mainly centered on the elderly who have opted to stay at home instead of home for the elderly. With more of the elderly population trying to stay at home for as long as

possible as opposed to going into nursing care, home health value purchasing has become an increasing necessity to help maintain the healthcare that they actually need

Assignment

    a. Discuss each of the factors affecting the demand for health insurance. Indicate the effect that each has on Demand.

    b. What is adverse selection? How does its existence affect the market for health insurance? What are some ways in which insurance companies try to protect themselves from adverse selection?

    c. How will a change in the price of medical care affect the demand for health insurance?

    d. How would you use utility analysis to analyze the following statement?

    e. Consumers should purchase health insurance policies that cover 100 percent of all medical care expenses. Anything less than 100 percent coverage reflects either irrational consumer behavior or market failures in the insurance industry.

References

Zweifel, P., Breyer F., Kifmann, M. (2009). *Health Economics*, 2nd ed. DOI10.1007/978-3-540-68540-1_1, Springer-Verlag Berlin Heidelberg.

Folland, S., Goodman, A. and Stano, M. (2016). The *Economics of Health and Health Care*. Pearson International Edition Seventh Edition. Routledge. ISBN-10: 1292020512 or ISBN-13: 978-1292020518

Rice, T.H. and Unruh, L. (2015). *The Economics of Health Reconsidered*. Chicago: Health Administration Press 4th Edition

Feldstein, P.J. (2012) *Health Care Economics*, 7th Edition, Cengage Learning Publishers Inc. ISBN-10: 1111313261 | ISBN-13: 9781111313265

Witter, S., Ensor, T., Jowett, M. and Thompson, R. (20015. *Health Economics for Developing Countries: A Practical Guide*. KIT Publishers, 2015. ISBN-13: 978-9460221316 ISBN-10: 9460221319

Jack, W(199). *Principle of health Economics for developing countries*. World Bank Development studies. ISBN 0-8213-4571-0

ACKNOWLEDGEMENT

The African Economic Research Consortium (AERC) wishes to acknowledge and express its immense gratitude to the following resource persons, for their tireless efforts and valuable contribution in the development and compilation of this teaching module and other associated learning materials.

1. Prof. M. Femi Ayadi, University of Houston-Clear Lake, Houston, Texas, USA. (Email: ayadim@uhcl.edu; mfayadi@gmail.com);

2. Prof. Bruno L. Yawe, Makerere University, Uganda. (Email: byawe@bams.mak.ac.ug; Byawe2010@gmail.com; byawe@yahoo.com);

3. Prof. Felix Masiye, University of Zambia, Zambia. (Email: felix.masiye@unza.zm; fmasiye@yahoo.com);

4. Dr. Moses K. Muriithi, University of Nairobi, Kenya. (Email: mkmuriithi@uonbi.ac.ke; Mmurithi13@gmail.com).

Thank you.
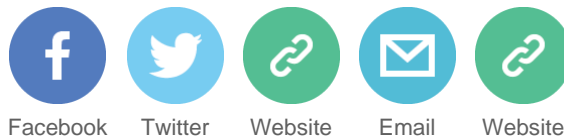
Facebook   Twitter   Website   Email   Website

Our mailing address is:
African Economic Research Consortium (AERC)
3rd Floor, Middle East Bank Towers, Jakaya Kikwete Road
P. O. Box 62882
00200 Nairobi
Kenya