



Supply-induced demand in the private health-care sector: Theoretical considerations

Mariné Erasmus and Helen Kean

Recently announced increases to medical scheme premiums were met with great surprise by many medical scheme members, who now have to absorb some of the highest premium increases of the past few years. Both the Government Employees Medical Scheme (GEMS) and Discovery Health have made statements in the popular press, suggesting that (amongst other factors) an increase in the supply of private hospital beds is to blame.

Given Econex's interest and previous research investigating the drivers of utilisation in the South African private healthcare market, this note approaches the concept of supply-induced demand (SID) from an academic perspective. The focus is on the relevant economic theory and theoretical considerations if one were to attempt measuring this phenomenon.

It is shown that a clear definition and explanation of how SID is hypothesised to operate in the local market are key to designing an econometric model to measure its effect on utilisation and costs. The economic literature further highlights the importance of correctly defining the market/ sample used for the study of SID, as well as the importance of identifying and controlling for the initial state of the market (i.e. whether demand and supply are in equilibrium to begin with) before measuring SID. Finally, we discuss that proving causation and including the relevant statistical controls in such a measurement exercise would be crucial for the final results obtained.

This research note therefore argues that, in measuring SID, a comprehensive understanding of the phenomenon and a robust econometric model are required, especially if results are to be used for policy formation.

About ECONEX

ECONEX is an economics consultancy that offers in-depth economic analysis, covering competition economics, international trade, strategic analysis and regulatory work. The company was co-founded by Prof Nicola Theron and Prof Rachel Jafta during 2005. Both these economists have a wealth of consulting experience in the fields of competition and trade economics. They also teach courses in competition economics and international trade at Stellenbosch University. For more information on our services, as well as the economists and academic associates working at and with Econex, visit our website at www.econex.co.za.

1 Introduction

Members of medical schemes in South Africa have recently been informed of higher than anticipated increases to their monthly premiums. The double-digit premium increases for 2017 are some of the highest that the industry has seen in the past couple of years. Many experts and industry players have offered explanations for these increases, few of which seem to be based on robust research.

Econex has previously analysed in detail the demand side factors driving utilisation increases and causing increased medical scheme expenditure on private hospitals specifically.¹ More recently however, the supply side drivers of utilisation have come under the looking glass. A presentation² by Dr Guni Goolab, the principal officer of the Government Employees Medical Scheme (GEMS), at the Board of Healthcare Funders' annual conference this year, states that increased supply equals increased utilisation. Particularly, "There is a correlation between increase in supply (especially hospital beds) and an increase in utilisation."

This concept, where an increase in supply is believed to cause

an increase in demand (utilisation), is termed 'supply-induced demand' (SID) by economists. Given Econex's interest in drivers of utilisation in the private healthcare sector in South Africa, as well as our academic interest in market dynamics such as this, we consider in this note the theoretical underpinnings for investigating SID as an explanatory variable for the high premium increases in the local market.

1.1 Relevant South African context

SID has for decades been of interest to researchers within the field of economics, and more specifically within the field of health economics. SID refers to the economic theory describing a phenomenon where the demand for a product or service is created *after* it is supplied; i.e. it is assumed that the supply of a product/ service 'induces' the demand for that particular product/ service. This is different from 'normal' market dynamics where most often there is a demand for a product/ service to which suppliers respond by producing (more of) that good.

As mentioned above, SID is of particular interest in the healthcare market as it is expected to explain some of the increases in

healthcare costs. Globally, it is generally accepted that healthcare costs rise faster than headline inflation. Within this context stakeholders and researchers are interested in understanding what portions of healthcare costs are driven by volume and price effects – and then at a deeper level, which portions of the volume effects are driven by demand and supply side effects. SID therefore comes into question in the study of the latter 'supply side effects'. For this reason a wealth of literature exists on the topic, with research and measurement attempts dating back to the 1930s. This disaggregation of effects is also of significant interest in the context of South African private healthcare, where utilisation (volumes) is widely acknowledged by stakeholders to be a key driver of increased medical scheme claims inflation.

Locally, and in addition to the concerns regarding high medical scheme premium increases, SID has been brought into question in the context of the Competition Commission's ('CC') local Health Market Inquiry ('HMI'), the aim of which is to better understand the drivers of costs within the private healthcare sector of South Africa. In particular, Discovery Health discusses that "One of the fun-

1. Econex Research Report 2 (September 2015), Research Note 36 (December 2014). Available at: www.econex.co.za

2. Available at: <http://www.bhfglobal.com/bhf-conference-2016/>

damental drivers that we are seeing quite recently and that, to be honest, is of great concern to us, is the impact of new private hospitals opening up recently... This is a very clear example of supply induced demand because our data shows that when new hospitals open in one particular geographic area often close to existing hospitals the level of, the water level of admissions just rises in total.”³ Similar to the comments from GEMS (mentioned above), this same argument is also put forward by Discovery Health in discussion of imminent premium increases.⁴

Interestingly, both GEMS and Discovery Health focus comments on SID as potentially ‘induced’ by hospitals. This is distinct to the wealth of economic literature which focuses on SID as potentially ‘induced’ by doctors. Logically, the latter follows on from the reasoning that doctors are generally the agents in a hospital; the hospital itself cannot ‘act’ to ‘induce’ care. Especially in the South African healthcare sector where private hospitals are prohibited from employing doctors, the focus on hospitals inducing demand is indeed strange.

Within the above broad context, this note draws on our

own experience, as well as a wealth of literature, to discuss what is (and is not) known and agreed upon regarding the phenomenon of SID. This provides a platform for further discussion on the topic and its applicability to the local (private) healthcare sector of South Africa.

1.2 Note structure

This note is set out as follows:

- Section 2 starts by discussing the economic concept of SID and the importance of understanding what exactly is being measured, if one were to attempt such an analysis.
- Section 3 goes on to explain certain aspects of this measurement process (i.e. an econometric analysis) that have been shown to be problematic. In particular, the various factors to consider when defining the market, how the initial conditions in the market may in fact explain the increased supply, and the need to establish causality in one’s assessment – these are each in turn explained. The need for a theoretically and practically sound mechanism to be put forward prior to such a detailed analysis, is again emphasised.
- Section 4 discusses the central role that the doctor plays in the utilisation of hospitals and the factors influencing doctors’ decisions.
- Section 5 explains that it is additionally important to consider how effective the ‘induced’ care is by analysing the change in health outcomes of patients prior and after receiving care, in order to establish whether the treatment was in fact necessary or not.
- Section 6 goes on to discuss the potential policy options to address increased utilisation.
- Section 7 concludes.

2 The economic concept of SID

2.1 Defining SID

In order to study the phenomenon of SID, one has to first be clear about the definition of SID that is in question. The complexity of this is illustrated by the relevant economic literature. Attempts to research and measure SID, specifically in the healthcare market, date back to the 1930s.⁵ However, despite decades of research, there is little agreement on the definition of SID, specifically applicable to the healthcare market.

3. Competition Commission Health Market Inquiry, 2016. Transcript of Hearing 3/ Day 2 held at Cape Town International Convention Centre on 2nd March 2016.

4. See e.g.: <http://www.businesslive.co.za/bd/companies/healthcare/2016-09-20-discovery-delivers-on-hefty-premiums-hike/>

5. Glover, J.A. (1938). The incidence of tonsillectomy in children. *Proceeding of the Royal Society of Medicine*, 31. p 1219 – 1236.

Chandra *et al.*⁶ for example explain the definition of SID to be “a physician providing care that a fully informed patient would not choose for himself”, while Leonard *et al.*⁷ explain SID as “the phenomenon of physicians deviating from their agency responsibilities to provide unnecessary care with the main objective of increasing their own pecuniary resources”, therefore introducing the concept of financial gains amongst doctors. Rossiter and Wilensky⁸ view a doctor initiating a consultation as SID, while Hay and Leahy⁹ explain that “the concept of physician-induced demand relies heavily on the notion of a knowledge imbalance between provider and consumer that the physician can and will exploit for financial gain”.

More broadly, Leape¹⁰ relies on the notion of ‘unnecessary surgery’ to indicate SID. He however explains the inadequacy of such a measure, explaining that it is often confused with ‘discretionary’ or ‘elective’ surgery, where the former is surgery that will not affect the physical functioning of the individual but is performed with the purpose of

increasing their sense of well-being.¹¹ ‘Elective’ surgery refers to surgery that does not have great risk of harming life or health of patients and may therefore be scheduled at any time. In parallel, ‘unnecessary’ may also indicate higher frequency of operations than what is considered normal. It also reflects patient preferences.¹¹

Other researchers further refer to ‘inappropriate’ surgery as surgery for which the negative consequences (namely mortality, morbidity, anxiety, pain and productive time lost) outweigh the positive effects (for example improved life expectancy, pain relief, improved physical functioning and reduced anxiety).¹² There is, however, very little consensus on how these criteria are to be established.¹²

From an economist’s perspective, SID may also be considered as the ability of doctors to increase patients’ demand for healthcare without lowering the price.

It is clear then that multiple definitions and measures of SID exist, focussing on different

characteristics and highlighting different areas of concern. In each of the above-mentioned cases, the measurement of SID will rely entirely on the definition considered.

Importantly, few (if any) definitions of SID refer explicitly to the role of hospitals or health-care facilities in their discussion of the phenomenon, with the focus rather on the doctor. This follows on from the fact that doctors are generally the agents in a hospital; the hospital itself cannot ‘act’ to ‘induce’ care.

2.2 The practical mechanism of SID studied

Alongside clarity on the economic definition of SID, the literature indicates that one should also be clear about the mechanism by which it practically works. In this regard a researcher should be able to simply explain the mechanism by which supply is hypothesised to be ‘induced’.

This should make sense both theoretically and practically, ahead of any technical testing; i.e. a study of SID should be

6. Chandra, A; Cutler, D & Song, Z. 2012. *Who ordered that? The economics of treatment choices in medical care. Handbook of Health Economics* 2:397-432. Page 14.7.

7. Leonard, C; Stordeur, S & Roberfroid, D. 2009. *Association between physician density and health care consumption: A systematic review of the evidence. Health Policy* 91(2):121-134. Page 121-122.

8. Rossiter, LF & Wilensky GR. 1981. *Identification of physician-induced demand. The Journal of Human Resources* 19(2):231-244.

9. Hay, J & Leahy, M. 1982. *Physician-induced demand: An empirical analysis of the Consumer Information Gap. Journal of Health Economics* 1(3):231-244.

10. Leape LL. 1989. *Unnecessary Surgery. Pew Memorial Trust Policy Synthesis: 7. Health Services Research* 24(3):351-407.

11. Leape LL. 1989. *Unnecessary Surgery. Pew Memorial Trust Policy Synthesis: 7. Health Services Research* 24(3):351-407. Page 353.

12. Leape LL. 1989. *Unnecessary Surgery. Pew Memorial Trust Policy Synthesis: 7. Health Services Research* 24(3):351-407. Page 380.

supported by a well-structured theory ahead of detailed testing.

Per illustration and in the local context, if one hypothesises SID as driven by hospitals, one would need to put forward a well-founded theory to support this. How the hospitals are able to act without the agency of doctors, and without constraint from medical schemes or their administrators, would need to be simply explained ahead of any detailed testing.

In the case that such logical explanation is absent, it would be extremely difficult to follow what exactly is in question or what mechanism is to be measured statistically.

3 Specifications of SID measurement

Notwithstanding differences in the economic concept of SID studied, the literature is clear that isolating and measuring SID in healthcare markets is a complex task.

Labelle *et al.*¹³ for example highlight some broad criticisms of SID studies, including:

- Lack of rigorous theoretical models: Results obtained in many SID models are consistent with both inducement models and neo-

classical models, meaning that results are ambiguous. Furthermore, models have been criticised as they are incomplete and in some cases rely on micro behavioural models that cannot be tested, such as the utility maximisation of doctors.

- Measurement error: It is found that SID models are often criticised regarding issues of measurement, resulting most frequently from making inferences about micro behaviour from aggregated data. Unrepresentative sampling is also often an issue.
- Specification error in econometric models: Omitted variable bias (in particular, health status) often characterises these models, as well as under-identification of healthcare demand equations. Models also often fail to acknowledge endogeneity amongst independent variables, such as the doctor-population ratio.

If one therefore attempts to measure SID econometrically, it is important to be explicit about the specification of the model used in order to avoid such issues. Below we draw on our own expertise and address some of the issues which we expect would need to be further considered.

3.1 Market definition

SID is generally studied in the context of a particular healthcare market, within which supply and demand are analysed. In order to carry out such an analysis one has to be clear about the market (or the sample) studied, particularly to understand whether trends indicative within that market are a function of some occurrence within that market itself (for example SID), or rather as a function of another market which has spill-over effects.

For example, one may observe healthcare supply trends increasing in a particular market. But this may be a benign function of decreased supply in other markets. Therefore, not capturing these effects may lead one to arrive at an incorrect finding of what is observed and the underlying reasons thereof.

Within the study of healthcare economics, a great deal of literature and techniques exist to assist with market definition – be it for the doctor or for healthcare facilities. These illustrate that one should be clear about the precise product and geographic market of study, in particular recognising that many specific healthcare services are not perfect substitutes for each other and that this interchangeability

13. Labelle, R; Stoddart, G & Rice, T. 1994. A re-examination of the meaning and importance of supplier-induced demand. *Journal of Health Economics*. *Journal of Health Economics* 13(3):347-368. Page 351.

also determines the geographic market.

In the healthcare sector, for example, it is reasonable to assume that patients would be willing to travel different distances to receive different types of treatment. Patients may be more willing to travel further for complex or life-threatening care than they would for relatively minor or routine healthcare. This assertion was confirmed in a seminal hospital merger¹⁴ in the United States, where the court found that separate markets for healthcare exist – one for primary and secondary healthcare services and a separate one for tertiary care services (e.g. cardiac surgeries or oncology therapies). In that case, each market defined in this way was identified as having a different geographic scope, with a more expansive geographic market defined for tertiary care than for primary and secondary care.

Other factors have also come into question in distinguishing healthcare markets. Chiswick¹⁵ discusses that differences in hospitals' social and physical amenities, distance from patients' homes and families, religious affiliations, and places where doctors have admission rights, are some of the factors that differentiate hospitals from each oth-

er.¹⁵ In the South African private healthcare sector, we expect that medical scheme provider networks – for both hospitals and doctors – would also play a role in defining healthcare markets. These networks strongly incentivise patients to seek care at particular hospitals and/ or with particular doctors (who often practise from hospitals), thereby driving patient flows.

Whilst the above discusses only some of the factors that will be relevant when one chooses a market and sample in which to study SID, it is additionally important to recognise that market definition may be dynamic – i.e. changing over time as demand and supply side factors adjust – for example as new facilities/ practitioners become available. The onus would therefore be on the researcher to illustrate how, if at all, this has been controlled for, particularly if the study period is of a reasonable time period and some variation in these factors is apparent.

3.2 Assumptions of initial conditions

Despite many differences in the economic concept of SID, as well as the market/ sample in which this phenomenon is studied, all such studies consider how utilisation/ demand adjusts

after a change in some supply side factor. **Importantly, such a model and related analysis requires an assumption of initial conditions – e.g. pertaining to whether the model does or does not assume supply and demand to be in equilibrium to begin with. Evidently, without such an assumption, it would be impossible to determine whether demand that correlates with an increase in supply is to be considered 'excessive' or 'induced'.**

For example, it is plausible that individuals in many regions experience pent-up healthcare demand, long travel times and long waiting times. In this case a (positive) supply shock via an increase in healthcare facilities or practitioners would bring such a region closer to equilibrium, where demand and supply are now in balance. However, in the absence of recognising an initial excess of demand over supply, this may (incorrectly) lead one to interpret that the increased demand that followed the increased supply is 'excessive'.

Per illustration from the economic theory, when the supply curve for a product or service shifts, it typically results in a new equilibrium between demand and supply: for an outwards shift of the supply curve, a larger quantity

14. *U.S. v. Long Island Jewish medical Center*. 983 F. Supp. 121 (1997).

15. Chiswick, BR. 1976. Regional variations in hospital occupancy rates. *Journal of Urban Economics* 3(2):113-126. Page 119.

of the product or service is sold, as the supply curve now intersects the existing demand curve at a new equilibrium point with higher quantities. Note however that there is a fundamental difference between a movement along the demand curve due to greater supply (which is what happens here) and an actual change in the demand function (where the demand curve would also shift). I.e. supply did not 'cause' greater demand, the existing demand was unsatisfied by existing supply previously.

Auster and Oaxaca¹⁶ similarly speak to the considerable care needed in testing the hypothesis that healthcare providers are able to shift the demand curve for their services. In the case of SID, the presumption is that suppliers (healthcare providers) are able to shift the demand curve outwards (i.e. increase demand). In other words, the shift in the demand curve follows a shift in the supply curve, with doctors wanting to increase their supply of services and using certain measures to induce a shift in demand. However, the neoclassical presumption is that suppliers will respond to excess demand because of supply being responsive to price. In other words, the demand curve shifts first, and the supply curve fol-

lows. In this case, increased supply and utilisation of healthcare services is not supply-induced demand, but rather a response by healthcare service providers to excess or 'pent up' demand.

Finally, Carlsen & Grytten¹⁷ argue that there are various alternative reasons that could explain increased utilisation associated with increased supply. The first is that increased utilisation reflects patient preferences and rationing of care prior to the increase in supply, rather than demand inducement. Due to less (implicit) rationing in the form of e.g. shorter waiting times and travel costs, patients may be more likely to access services after the increase in supply. It is therefore explained that the state of the market before the increase of the supply is indicative of a certain level of rationing of healthcare resources. As more resources become available, less rationing is required and in addition, people who are less sick can also receive the required services.

Another reason put forward by Carlsen & Grytten is that physician density is endogenous – where a doctor decides to start a practice depends on market conditions (as doctors are expected to migrate to areas

where there is excess demand). Based on the above discussion, it is concluded that one of the most problematic factors in studying SID is the 'state' of the market before the increase in supply, albeit hospital beds, doctors or both, that are increased. This challenge is well described by Chandra *et al.*: "Supply may be high because demand is high. Distinguishing whether greater equilibrium quantities are provided because supply is high or demand is high is always difficult, and is virtually impossible when prices are not allowed to adjust."¹⁸

Given this, any model that aims to use observed variables of final quantities of healthcare supply, to make some claim about unobservable demand or its drivers, must solve this identification problem of initial conditions. One cannot, based on a correlation between equilibrium quantities and a supply shift, simply conclude that demand is being influenced or so-called 'induced' by supply.

The above is particularly relevant in the South African healthcare sector, where we know that there is a lack of overall resources and hence some degree of pent-up demand. We also know that there are bar-

16. Auster, RD & Oaxaca RL. 1981. Identification of supplier induced demand in the health care sector. *The Journal of Human Resources* 16(3):327-340. Page 328.

17. Carlsen, F & Grytten, J. 1998. More physicians: Improved availability or induced demand? *Health Economics* 7(6):495-508. Page 495-496.17.

18. Chandra, A; Cutler, D & Song, Z. 2012. Who ordered that? *The economics of treatment choices in medical care. Handbook of Health Economics* 2: 397-432. Page 415.

riers to entry in the healthcare market (for example hospital licenses that have historically taken significant amounts of time to be granted, as well as limitations on training of new doctors) such that healthcare supply has not always been able to directly and timeously respond to demand. Within this context, any study of SID would need to carefully establish the 'state' of the market before the increase in supply and any observable demand changes that follow.

3.3 Relevant controls and proof of causality

In addition to the factors already discussed and pertaining to the considerations one should take into account in studying SID, the literature also shows that one would need to carefully control for all other reasonable determinants of demand when studying the effect of a specified supply factor on demand. I.e. one would also need to prove causality of any relationship between supply and demand by utilising relevant measurement methods, as well as including all relevant controls.

Importantly, economic theory illustrates that statistical correlation does not prove causa-

tion. Therefore studying correlations between – in the case of SID – demand and supply does not per se establish a relationship between these factors. In order to prove causality one would, at the least, need to control for other confounding factors.

For example, the supply of practitioners, the supply of facilities, the supply of medical technologies, and socio-economic and demographic factors all impact on healthcare demand levels. If the effect of one of the supply factors on demand is to be isolated, at the least all of the other factors mentioned (as well as any others relevant to the sample studied) must be controlled for. One cannot, based on a simple correlation between a supply variable and a demand variable, conclude that that demand is being influenced by that specific supply factor.

The issue of causality is however not one of statistical robustness alone. Notwithstanding the ideal requirement for statistical robustness, economists often have to rely on correlations of the above form to study demand behaviour. In those instances the analysis must in-

clude either econometric attempts at achieving identification, or qualitative attempts at achieving the same: there must be a well-founded theory that explicitly and exclusively links demand to that particular driver. In other words (and as mentioned in section 2.2) one must have a well-founded theory of how SID practically takes place.

As such, if one were to measure SID in the South African private healthcare market, a logical explanation of how increased supply is expected to induce increased demand must be put forth. If the hypothesis is that an increase in the supply of hospital beds 'causes' an increase in the demand for hospital services, that mechanism should be explained adequately.

It is difficult to understand how a new hospital would be able to admit patients without any doctor involved in that process. In the South African environment, admission decisions are made by doctors (independent of hospitals), with some constraints imposed via managed care organisations and medical scheme administrators. *Hospitals* are not able to induce patients into the facilities, nor are these hospitals able to keep patients there once admitted.

More Information

ECONEX regularly publishes Research Notes on various relevant issues in South African competition, trade and applied economics. For access to previous editions of Research Notes, or other research reports and published articles, go to: www.econex.co.za. If you want to add your name to our mailing list, please send an e-mail to iris@econex.co.za

4 The central role of the Doctor

In line with the concerns discussed above, the one factor that there is some consensus on in the study of SID, is that it is largely understood to be the doctor who would – if at all – have the ability to ‘induce’ demand, and not the facility or institutional setting where he/she delivers the service. While there are papers exploring the role of the hospital or other service being made available, it is generally the case that these services cannot be utilised without a doctor recommending it.

In the economic literature, doctors per capita have been found to be a more powerful predictor of utilisation than bed supply. A study by Lewis¹⁹ in fact found that almost half of the variation observed in healthcare utilisation could be explained by the increased availability of doctors. Cross-country comparisons have also shown that higher rates of surgery per capita are matched by higher numbers of surgeons (Bunker²⁰ for example

explores this phenomenon in the US and Great Britain). Similar results are found by Vayla²¹ when comparing Canada with England and Wales. This result is consistently found: populations with a greater number of surgeons have more operations.²²

Following on from the above literature, there also exists a large body of literature exploring the doctor’s decision-making function, i.e. how, why and when he/ she decides to treat a patient, as well as what that treatment entails. The availability of e.g. hospital beds certainly influences the doctors decision to admit a patient, but “it is difficult to believe that many physicians recommend an operation simply because an empty bed is available – or that they fail to do so because a patient might have to wait a week to get into a hospital”.²²

Evidence of differences in the practise style or ‘philosophy’ of doctors across regions have been shown to impact on the rate of healthcare utilisation. Considerable uncertainty around medical decisions means that there is diversity in doctors’ willingness

to innovate and experiment. It often takes younger, more innovative doctors in an area for an extended period of time before the behaviour of all doctors in the region changes. This often results in significant differences in utilisation across regions – regularly accepted as an indication of SID. Higher levels of uncertainty result in more regional variation, while procedures with lower levels of uncertainty tend to vary less across regions.

Numerous factors influence doctors’ decision-making and it is unlikely that all but a few doctors would knowingly perform unnecessary surgery.²³ Another such factor is the institution at which they did their training. Different hospitals are known to employ different methods for treating the same conditions, and often whether a specialist employs a particular treatment is predicted by whether or not their colleague or supervisor employed the same treatment.²⁴

Despite the above studies, the ‘link’ between the doctor and hospital is nevertheless important. Kroneman and Siegers²⁵ expressly model the relationship

19. Lewis, CE. 1969. *Variations in the Incidence of Surgery*. *New England Journal of Medicine* 281(16):880-884.

20. Bunker, JP. 1970. *Surgical Manpower. A Comparison of Operations and Surgeons in the United States and in England and Wales*. *Journal of Medicine* 282(3):135-144.

21. Vayda, E. 1973. *A Comparison of Surgical Rates in Canada and in England and Wales*. *Manpower*. *The New England Journal of Medicine* 289(23): 1224-1229.

22. Leape LL. 1989. *Unnecessary Surgery*. *Pew Memorial Trust Policy Synthesis: 7. Health Services Research* 24(3):351-407. Page 360.

23. Leape LL. 1989. *Unnecessary Surgery*. *Pew Memorial Trust Policy Synthesis: 7. Health Services Research* 24(3):351-407. Page 363, 386.

24. Chandra, A; Cutler, D & Song, Z. 2012. *Who ordered that? The economics of treatment choices in medical care*. *Handbook of Health Economics* 2: 397-432. Page 418.

25. Kroneman, M & Siegers, JJ. 2004. *The effect of hospital bed reduction on the use of beds: A comparative study of 10 European countries*. *Social Science & Medicine* 59(8):1731-1740. Page 1733.

between doctors and hospitals in the phenomenon of SID, pointing out that differences in the incentives created by this relationship may have opposite effects on utilisation. They explain that alignment of incentives of doctors and hospitals is an essential determinant of SID and requires careful consideration in any attempt to measure it.²⁶

Importantly in the local context, since doctors are prohibited from being employed by private hospitals, hospitals (and hospital managers) have only limited mechanisms to 'control' the doctors' behaviour. And while it is argued that hospitals are not able to incentivise doctors to admit patients unnecessarily, it should be noted that there are many control mechanisms aimed at curtailing any potential excessive admission of patients by doctors, including clinical committees and, most importantly, managed care by medical schemes.

More specifically, the regulatory framework within which doctors operate include ethical rules and codes of conduct that would not easily allow for unnecessary surgery that is not in the patient's best interest. In addition to these constraints,

the managed care environment requires a patient to obtain authorisation from the particular medical scheme prior to admission; and includes strict rules regarding the length of stay in hospital. A hypothetical increase in unnecessary admissions would therefore also bring into question the effectiveness of managed care organisations.

5 Health outcomes: Cost versus Benefits

Finally, and of central importance to the study of SID, the relevant economic literature also highlights that without considering the health outcomes of the services received by patients, one cannot determine whether those services were in fact necessary or not.

What therefore becomes important in any analysis of SID, is the health status of patients pre and post their utilisation of medical services. Without considering the health outcomes of the services received, one cannot define those services as unnecessary. Patients may be better off, implying that the services provided were likely a normal market response to existing demand. A proper assessment of

SID would therefore include a type of cost-benefit analysis to determine whether the services delivered were indeed unnecessary.

However, in the vast literature on the topic of SID, little attention has been paid to how doctor-induced demand affects the health outcomes of patients. Any consideration of how the behaviour of doctors impacts the health outcomes of patients seems to have been lost.²⁷ This is important since researchers argue for analytically distinct types of SID which have different implications for health outcomes. A thorough investigation of SID therefore requires that the health outcomes of patients resulting from this increased demand be considered.²⁸

Per illustration, and in the context of South African private healthcare, it would be relevant to consider the effect of improved health outcomes on medical schemes' longer term finances; and the effect of increased supply on competition, prices, and hence medical scheme payments. Focussing solely on immediate cost implications of increased demand and supply would otherwise not be particularly insightful.

26. Kroneman, M & Siegers, JJ. 2004. *The effect of hospital bed reduction on the use of beds: A comparative study of 10 European countries. Social Science & Medicine* 59(8):1731-1740. Page 1733.

27. Labelle, R; Stoddart, G & Rice, T. 1994. *A re-examination of the meaning and importance of supplier-induced demand. Journal of Health Economics. Journal of Health Economics* 13(3):347-368. Page 349.

28. Labelle, R; Stoddart, G & Rice, T. 1994. *A re-examination of the meaning and importance of supplier-induced demand. Journal of Health Economics. Journal of Health Economics* 13(3):347-368. Page 351.

In summary then, cost containment seems to be the main issue at hand in the SID literature, at the expense of the impact on health status, distributional implications, and the net social benefits that may result from SID (if indeed proven).²⁹

Important to consider in understanding SID and largely absent in the literature, is how effective this 'induced' care is.³⁰

The fact that the doctor often acts or decides on behalf of the patient that extra care should be provided, but whether the additional care is effective and improves the well-being of the patient, should also be considered.

6 Policy responses

After reviewing many academic papers on SID, it is found that the suggested policy response is almost always a change in the reimbursement structure of doctors, as well as utilisation review or the introduction of managed care. It is suggested that these methods are more effective in controlling demand than supply-side regulations would be.

Especially in the South African private sector environment where managed care organisations are expected to play

an important role, one should question how any patient can be admitted without the clinical need to be admitted and obtaining 'pre-authorisation' from the medical scheme or administrator.

If expenditure control was the only objective to be pursued by policy-makers, then certain regulatory strategies may be effective. However, in order to also achieve allocative efficiency, more needs to be done to align the private incentives of doctors with that of social valuations. This could be achieved by, for example adjusting doctor reimbursement to compel doctors to consider the total cost of the service provided rather than just the portion paid by individual patients. Another possibility is to adjust prices to reflect the marginal social cost of providing services, although this is what health insurance systems try to avoid by ensuring that individuals have access to healthcare services regardless of their ability to pay.³¹

7 Concluding remarks

Against the backdrop of high medical scheme premium increases, this research note set out to explore the theoretical ar-

guments potentially supporting the theory of SID as put forth by GEMS and Discovery Health. These two medical schemes (respectively the largest restricted and open schemes) have commented that premium increases are, amongst others, driven by utilisation increases that are in turn caused by an increase in the supply of private hospital beds in South Africa.

Given Econex's previous interest in the drivers of utilisation in the private healthcare market, we considered (from an academic perspective) the relevance of SID, as well as how one would go about measuring this phenomenon statistically. The relevant economic literature showed that a clear definition of SID prior to any analysis, as well as a logical description of how SID is expected to operate within the South African market, is key to approaching the research question.

From the discussion, it was further understood that an econometric analysis of this nature would indeed be challenging, given the significant data requirements and statistical controls required. Our expert knowledge of defining healthcare markets and designing econometric models, support-

29. Labelle, R; Stoddart, G & Rice, T. 1994. A re-examination of the meaning and importance of supplier-induced demand. *Journal of Health Economics*. *Journal of Health Economics* 13(3):347-368. Page 350.

30. Labelle, R; Stoddart, G & Rice, T. 1994. A re-examination of the meaning and importance of supplier-induced demand. *Journal of Health Economics*. *Journal of Health Economics* 13(3):347-368. Page 354.

31. Labelle, R; Stoddart, G & Rice, T. 1994. A re-examination of the meaning and importance of supplier-induced demand. *Journal of Health Economics*. *Journal of Health Economics* 13(3):347-368. Page 362.

ed by evidence from the relevant literature, proved that the specification of such a model attempting to isolate the effect of SID in the healthcare market is crucially important to obtaining reliable results. Specifically the market definition, assumptions of initial conditions (state of the market ex ante), as well as proof of causality would be determining factors.

It was additionally pointed out that, contrary to the literature, the role of the doctor is amiss from the arguments regarding SID in the South African private healthcare market. Also, from high level remarks, it seems that consideration of improved health outcomes and the implications thereof have not been

analysed. This is a notoriously difficult factor to account for, but key to understanding the potential costs and benefits of SID as hypothesised.

Looking forward, an improved understanding of the drivers of utilisation (both demand and supply side factors) is imperative for ensuring a sustainable future for medical schemes in South Africa. As such, it is important that statistical models are correctly defined to critically assess these factors; taking account of the underlying literature and previous attempts to do so. Specifically when informing policy responses or regulatory changes aimed at cost containment, robust research must be performed.