

TOWARD A THEORY OF MARKET FAILURE IN THE SOUTH AFRICAN PRIVATE HEALTH SYSTEM

Alex M van den Heever

alex.vandenheever@wits.ac.za

Chair of Social Security Policy Management and Administration, Wits School of
Governance, University of the Witwatersrand, Johannesburg, South Africa

DRAFT REPORT

Do not distribute or quote without the permission of the author

March 2015

CONTENTS

CONTENTS.....	i
TABLES.....	ii
FIGURES.....	ii
Introduction.....	1
Methods.....	2
Health market failure in theory and practice.....	2
Theoretical case	2
Funders – a review	4
Healthcare Providers	5
Overview of the South African private health market	10
Overview.....	10
Medical schemes regulation to 1967.....	10
Tariffs and Payments: 1968 to 1986	11
The “Free Market” Reforms: 1984 to 1989	12
The period 1990 to 1994	14
Government medical scheme	14
Removal of mandatory benefits from medical schemes	15
The period of scheme deregulation 1994 to 2000.....	15
The period of medical re-regulation from 2000 to the present	21
Conclusions.....	29
Explaining hospital and specialist cost increases over the past two decades.....	30
The theory of market failure in context.....	34
Medical schemes	34
Hospitals and specialists	35
General practitioners.....	36

Other healthcare products	36
Resolving market failure in the South African private health system.....	37
Concluding remarks	40
References.....	41
Annexure 1: Private doctor and hospital response to schemes' attempts to introduce selective contracting in 1997	44

TABLES

Table 1: Private and public hospitals and bed estimates from 1976 to 2010	14
Table 2: Changes in national hospital market concentration based on acute beds, from 1996 to 2010	21
Table 3: Medical scheme beneficiaries and private hospital bed to population ratios for 2001 to 2010	33

FIGURES

Figure 1: Beneficiary changes from 1990 to 2010, distinguishing open from closed (restricted membership) schemes	16
Figure 2: Non health expenditure in medical schemes from 1974 to 2010 expressed as a percentage of Gross Contribution Income.....	17
Figure 3: Medical scheme gross contribution income (GCI) per average beneficiary per annum, 2010 prices (Rands) for the period 1990 to 2010.....	24
Figure 4: Medical scheme real per capita claims costs from 1981 to 2010 (2010 prices).....	26
Figure 5: Medical scheme real claim cost changes from 1993 to 2010 (pabpa) compared to GDP expressed on a common index with 1993 = 100	28
Figure 6: Index changes in real medical schemes expenditure on hospital and non-hospital claims compared to GDP 1994-2010 (1993 = 100).....	28

Introduction

Health systems are complex due to the large number of funding mechanisms, prioritisation mechanisms, role-players, vested interests, and government interventions. The wide variety of government intervention adds to the difficulties in understanding how health care markets work, and how they respond to different signals.

Some countries have large government funded and provided systems while others have market-based systems with a range of regulatory configurations. High-income countries also differ markedly from middle- and upper-middle income countries in the depth, uniformity, and costs of coverage. The absence of any ideal or best-practice model requires that individual countries develop strategies that make sense in context, while drawing on international case studies able to provide usable generalizable evidence.

Competition in the South African private health system provides a useful window into the opportunities for optimising private health markets, or conversely, the consequences of failing to do so. The size and relative maturity of the private health system offers the possibility of insights not necessarily available in other developing or even high-income country contexts.

This paper makes use of generally available evidence to build a theoretical approach that best explains market failure in the South African health market. From this vantage point an intervention framework is developed to counter the failures and optimise the performance of the market. The guiding premise of the article is the view that certain markets, of which healthcare is one, cannot perform optimally in the absence of an engineered framework that structurally corrects for market imperfections.

The paper therefore develops a set of tentative theories, using secondary data sources and qualitative insights that, if proven, would isolate the structural causes of market failure. It is not the purpose of this article to prove the theories, but rather to establish a *prima facie* case for their statement as a probable theory of market failure.

This paper questions the view that health market failure is fatal, and can only be addressed through substitutive government-run insurance or delivery and/or administered price systems. Instead, government's role, if focused on the maintenance of the system-wide framework within which competition occurs, would be able to drive social outcomes not otherwise possible through either the unaided market or government.

Methods

The approach in this paper is to draw together secondary information from various sources to frame a plausible theory of market failure specific to the South African context. The purpose is not to prove the theory, but rather to set out a case for a narrowed set of parameters for the prioritisation of future research.

The results of this article take the form of an over-riding theory of market failure in South Africa that needs to be tested to confirm, qualify or disprove the theory. The analysis leading to the theoretical framework combines a review of conventional theories of health market failure with an examination of domestic market trends and outcomes.

Health market failure in theory and practice

Theoretical case

Healthcare markets are argued to suffer from failures on both the demand and supply side that can both harm access to coverage and drive up costs beyond what a properly functioning market would permit. The interaction between insured markets, which emerge naturally due to the inability of spot markets for health goods and services to deal with catastrophic expenses (for instance through loans), and supplier incentives, drive up costs. These include:

On the funding side:

- Information asymmetries between an insurer and the insured:
 - An insurance applicant is typically unable to understand the quality of the insurance sold to them;

- Insurance applicants, who have greater knowledge of their own health status than insurers, have an incentive to apply for insurance only when sick and leave after treatment (referred to as anti-selection); and
- Insured individuals facing no or limited point-of-service costs have an incentive to over-consume health care goods and services (or ex-post moral hazard).

On the supply side:

- Patients cannot shop around for services and compare prices and service quality, i.e. the search costs are high, and are vulnerable to supplier generated demand which can arise for two reasons:
 - Financial incentives of healthcare service providers who are able to benefit from advice they provide to structurally uninformed consumers; and
 - Professional incentives, whereby treating doctors seek to assist each patient to the maximum, without consideration of cost.

Consumers consequently face fundamental information asymmetries which structurally impact on their ability to make rational purchasing choices in their own interests:

- They do not know what to buy as the market, whether for insurance or health goods and services, is too complex for them to understand in the time they reasonably have to make a decision; and
- In the case of healthcare, they do not know what they are purchasing, and do not even know the value of what they've purchased after receiving it.

Health insurance providers respond to moral hazard from insurance applicants by risk-rating and risk-selection (refusing or increasing the cost of coverage for poor health risks). Healthcare providers, conversely, respond by providing more services, introducing technology at inefficient levels and costs, and charging more over time. This conduct does not need to be explicitly malicious to be damaging and can emerge purely from the accumulation of many decentralized decisions responding to a combination of professional and economic incentives.

According to Gronewald and Stiglitz (1986), whenever markets exhibit significant information asymmetries, involving moral hazard, adverse/anti-selection, or incomplete risk markets, the economy will not be efficient. These characteristics essentially define health markets and suggest they deviate substantially from the conditions required for a competitive market (Rothschild and Stiglitz, 1976).

Funders – a review

Health systems are diverse, with significant differences in the organization of funding and provision. Many health systems rely on non-competing funding arrangements irrespective of whether they involve private funders and providers, as in Germany, the Netherlands, and Belgium or public funders, such as the United Kingdom (UK), Canada, South Korea, and Taiwan.

Whereas the public funding models (such as the UK) often rely on global budgets as a form of expenditure constraint, multi-funder environments indirectly limit expenditure through controlling contribution increases and regulating prices (Germany and the Netherlands). South Korea and Taiwan, despite having a single purchaser, fund private services on a fee-for-service basis. This limits their ability to control costs on a global budget basis. They consequently ration indirectly through the application of very high co-payments.

The United States (US) is an example of competing funders of private services, at least for those families falling outside their two social insurance arrangements, Medicare (for those over the age of 65) and Medicaid (for those falling below a means test). However, each state regulates insurance differently. While some states prohibit risk-rating and risk-selection, others allow it. No states have mandatory minimum benefits and consequently product competition is extensive, making it next to impossible for consumers to prospectively determine the competing value-for-money of different offerings. There are also substantial access differences, depending upon whether you join an insurer through a group arrangement (such as an employer) or as an individual (when self-employed or working for a small employer).

Although the US provides interesting case studies for an analysis of insurer competition, its market is so weakly regulated that it's hard to exclude confounding factors in empirical analyses. For instance, price increases and quality reductions

resulting from provider concentration may be masked by competition failures on the funder side.

The evolution of so-called managed care in the US was at one point regarded as a market-based solution to rising costs. However, there appears to have been no systemic effects on costs or service quality despite geographically confined instances of success (Melnick et al, 1992). Nevertheless, some of the innovations to manage provider costs have been adopted more effectively by countries with far greater regulation on the funding side. For instance, the near universal acceptance of the Diagnostic Related Group (DRG) approach to reimbursing hospitals has been adopted in most OECD countries irrespective of whether they are funding public or private providers.

In all instances where global budgets and regulated contribution increases are applied, governments are compelled to negotiate the consequences with providers in one way or another. These negotiations can include the rationing of supply, and supply prices. On the whole most mature health systems (apart from the US) do not rely on competing funders to manage health costs as they appear to systemically lack the ability to do so, even when the incentives are present. However, no clear examples exist of countries that have tried to establish the conditions for such competition to work. Most experiments focus on making providers rather than funders compete.

Within developing countries the experience of multi-fund arrangements is as mixed as the performance of state systems, with very few adopting a regulatory regime sufficient to deal with the fairly well-understood market failures. For instance Chile established a system which provided for choice between public and private funders and providers. All workers must pay at least 7% of their income to either a state insurance scheme, which has dedicated providers, or one of 15 private schemes. There is, however, no regulation to prevent discrimination based on health status.

Healthcare Providers

Despite the interest in market-based solutions to both public and privately funded markets, the empirical evidence on the effects of competition on provider performance is ambiguous, due in part to difficulties in establishing causality. (Cooper, 2012).

Recent reforms in the English part of the National Health Service (NHS) do provide some opportunity to examine these effects. In 2006 publicly funded patients were permitted to choose their hospitals, with funding following the patient.

Cooper (2012) concludes that the introduction of provider competition in the English NHS has been a success, leading to reduced death rates, improvements in hospital quality, management and productivity with no harm to equity. The factors contributing to the success are seen as the combination of centralization and decentralization.

When motivating for the reforms the government argued that "*hospital competition in a market with fixed prices would catalyze providers to become more efficient, more responsive to patients and improve the quality of care hospitals delivered.*" (Cooper, 2012, p.7)

For these reforms to work the following conditions had to be satisfied (Cooper, 2012, pp.9-12):

- "*Hospital staff, including senior management, must be responsive to financial incentives*" which was achieved by having autonomous public hospitals accountable for financial viability, compete for patients.
- "*Patients must be interested in making choices*" which becomes important when local hospitals fail to address queues or quality. "*Evidence from the London Patient Choice Pilots suggests that hospital cleanliness and hospital waiting times were the two most influential factors over whether or not patients went to a non-local provider...*" (Cooper, 2012, p.11)
- "*Patients must have alternative providers that they can access*" which was achieved by allowing private providers to offer services to public patients.
- "*There must be information in the market to inform patients' and purchasers' choices*" which was achieved by providing public information on services, using GPs to act as agents of care (affecting 40% of patients), while own experiences and advice from friends and family were also important.
- "*Patients must be responsive to quality signals*". It was found that sicker (older patients and those with co-morbidities) and lower-income patients were very responsive to quality and waiting times respectively after the reforms.

Cooper (2012) stresses that the NHS reforms reflect that competition works when supported by concurrent policies which are often not present in other settings (such as the US). Importantly, studies conducted in the initial phases of payer-driven provider competition in the UK indicate a drop in quality, most probably due to the absence of quality signals. (Propper et al, 2004, p.1267). This suggests strongly that competition occurs principally on factors that are visible, and where consumers are willing and able to make alternative choices based on these factors. Whatever remains hidden cannot influence decisions.

"This suggests that it may have been a mistake to delay the publication of quality signal until some 10 years after the introduction of a market meant to rely on them." (Propper et al, 2004, p.1267)

Whereas the NHS experience reveals that providers will respond positively to competition on non-price factors, as prices are regulated centrally, an important question is whether provider price competition reveals positive results.

An important distinction exists between NHS public hospitals and their private counterparts. Although the NHS hospitals are publicly owned, they are managed as autonomous entities and made to compete against each other. Amalgamated (or merged) private hospitals, by way of contrast, cease to compete. Mergers nevertheless do offer the possibility of efficiencies through scale economies. There is however no evidence that mergers produce cost savings from increased economies that are passed on to patients. (Gaynor, 2012)

An important mechanism to achieve hospital and provider competition is selective contracting as noted by Gaynor (2012, p.22):

"Selective contracting by private payers is a mechanism which helps to create an environment where competition among providers is possible. If payers contract with every provider in the market, then they have little bargaining power. Policies that enhance the ability of payers to selectively contract with providers are important, but they will only be effective if there is sufficient choice among providers."

Gaynor (2012, p.15) also reaches some clear conclusions on hospital competition:

- Hospital competition leads to lower prices and quality improvements when prices are fixed administratively.
- No evidence exists of efficiencies arising due to mergers.
- Quality improvements arise from competition on quality, where patients are independently advised on which services to use (i.e. where they are enabled to make choices in their own interests).
- Price competition reduces prices provided hospitals have strong incentives to attract demand based on price concessions to responsive consumers (or party acting on their behalf).

Arguments opposing the standard economic view, that prices are lower and quality better in less concentrated markets, have been proposed in US-based merger cases, where the courts accepted evidence that more concentrated markets lower prices in the case of hospitals. (Dranove & Shanley, 1993). There is however a growing body of evidence that suggests that supports the conventional view.

Gaynor (2012, p.13) concludes, after reviewing a large number of studies from the US, that *"hospital consolidation raises prices, and by nontrivial amounts. Consolidated hospitals that are able to charge higher prices due to enhanced market power are able to do on an ongoing basis, making this a permanent rather than a transitory problem."*

Vogt & Town (2006) in a survey of event studies of hospital mergers in the US find that hospital process rose by 10% and above. When considering the affect of a merger on rivals, who also benefit from an increase in concentration, an increase of 40% was found in recent strong event study. Overall they find that hospital mergers in the 1990s increased prices by at least 5% *"and likely by significantly more"*.

Dranove & Shanley (1993) also argue that healthcare price competition effects must be evaluated in context. A distinction is made between "patient-driven" and "payer-driven" markets, with the latter able to mitigate the market failure of the former. Mergers occurring in patient-driven markets consequently have different effects to payer-driven.

Payer-driven markets exist where the funder (health insurer and purchaser) take direct responsibility for provider arrangements with consumers choosing the funder/insurer rather than the provider. Consumers choose funder/insurers based on the cost and quality of the health care providers.

Patient-driven health insurance markets are characterized by *passive* insurers focusing narrowly on the reimbursement of medical expenses incurred. Consumers choose funders/insurers simply on the cost of the insurance and the benefits covered. Patient-driven insurance markets can also exist with central state purchasers (as in South Korea and Taiwan) where simple fee-for-service reimbursement, a characteristic of patient-driven markets, is prevalent.

Market structure has furthermore been raised in the emerging private hospital and insurance markets in the UK with similar conclusions emerging from the Office of Fair Trading (OFT, 2011) to US-based studies.

Melnick et al (1992, pp.229-232) supporting the market structure findings of Cooper (2012), Gaynor (2012), and Dranove et al (1993) found similar results in a study of network contracts with the large not-for-profit private health insurer Blue-Cross.

"These results show that actual hospital prices are lower in more competitive markets. The policies encouraging selective contracting [reflecting a payer-driven market] have proven effective in transforming the structure of the private insurance industry and, in turn, producing changes in the nature of hospital competition. ... The study also shows that this pressure is fragile and in need of protection. Hospitals tend to operate in markets that are relatively concentrated at least from the perspective of the overall economy. A relatively small number of mergers in key markets would shift bargaining power and lead to increased prices."

Consistent with findings that concentration increases prices, Krishnan (2001) finds that DRG (diagnostic related grouper) prices, often regarded as a way to internalise over-supply incentives) are also higher in merging compared to non-merging hospitals. This suggests that DRGs as a method of contracting are also influenced by hospital market power and on their own do not resolve by focusing exclusively on moral hazard concerns.

Overview of the South African private health market

Overview

South Africa is characterised by two distinct systems of financing and providing health care. A general tax-funded system of public care financed and rendered by provincial governments which spends approximately 3.9% of Gross Domestic Product (GDP) in 2013/14 (National Treasury, 2012) and implicitly covers around 45 million people; and a contributory medical scheme funded system which spends around 3.1% of GDP (based on National Treasury, 2012) and explicitly covers around 8.8 million people (Council for Medical Schemes, 2014). Overall health expenditure is around 8.0% of GDP (National Treasury, 2012) when out-of-pocket expenditure and some minor social insurance arrangements are taken into account.

Medical schemes regulation to 1967

The private sector evolved largely in the post-war period in a policy vacuum, with government policy focused on general tax-funded health services for the non-paying population. As higher income earners had to pay for both public and private sector services, what would ultimately be framed in legislation as medical schemes became an important parallel funding vehicle to the state. As medical schemes pre-dated the development of the public health system (the first scheme was created in 1888), their evolution as a coherent system of funding healthcare emerged as a response to the need for insurance in the absence of access to free public services for income earners and their dependents.

By 1940 around 48 employer-based (or occupational) medical schemes existed. These increased significantly in number after the Second World War necessitating the development of some form of coordination. A supervisory structure, the Advisory Council for Medical Fund Societies, was consequently established in 1950. No regulatory framework was established at that time, and the Council's role was largely to represent the interests of schemes in negotiations with the Medical Association of South Africa (SAMA) on fees. (Department of Health, 2002).

Regulation of medical schemes occurred for the first time through the Friendly Societies Act implemented in 1956. Regulation was limited to registration and some minor prudential requirements. Consistent with the findings of two commissions of

inquiry, the need for an expanded regulatory framework was accepted and implemented in 1967 through the Medical Schemes Act No. 72 of 1967 which took over the functions of medical scheme regulation from the Friendly Societies Act. (Department of Health, 2002).

The intentions of the Medical Schemes Act were (Hansard, 1967, from Department of Health, 2002, p.21-22):

- *“To invent an insurance type of scheme to distribute the costs of medical expenses over a period of years”;*
- *“To retain the costs of medical expenses at a low level”;* and
- *“To co-ordinate and control the functioning of medical benefit and medical aid funds and to develop and propagate these schemes”.*

Tariffs and Payments: 1968 to 1986

The period from 1968 into the mid-1980s medical schemes predominantly reimbursed private professional fees and tariffs charged by public hospitals, promulgated in provincial ordinances, to people falling outside a means test. Professional fees for specialists and general practitioners therefore represented the most important expense of schemes, and annual negotiations were consequently quite heated. As the reimbursement of medical expenses represented the benefits offered by a scheme, the annual centrally determined tariff schedule was referred to as a “scale of benefits” (SOB).

As the SOB reflected the reimbursement schedule of all medical schemes, disagreement over annual changes resulted in some professionals choosing to opt out of the SOB. Where health professionals opted out and charged in excess of the SOB schemes reimbursed only up to the benefit level leaving the residual for the account of the member. As claims were managed through paper-based invoices, schemes adjusted to the opt out (contracted out) doctors by guaranteeing reimbursement only where invoices were consistent with the SOB (contracted in). A version of this continues to today.

A Remuneration Committee was set up in terms of the Medical Schemes Amendment Act, No. 95 of 1969, to investigate the tariff of fees at least every two years. This amendment sought to establish an arbitration mechanism to avoid further doctors

opting out of the scale of benefits as this was regarded as damaging to doctor/patient relationships. (Department of Health, 2002).

The Remuneration Committee was however ultimately not supported by doctors. It was alleged that the Act was seeking to control the medical profession and that the framework prejudiced the interests of doctors and dentists. By 1978 the Dental Society and the Medical Association were no longer prepared to cooperate with the Remuneration Committee. Consideration was then given to limiting the free choice of doctors to contract out. Publication of draft legislation to this effect however resulted in a further 1,600 medical practitioners contracting out. By this time 3,941 out of 14,000 medical practitioners had already contracted out. (Department of Health, 2002).

To resolve this conflict the Amendment Act No. 51 of 1978 abolished the Remuneration Committee and the Commission that made recommendations to the Council on fees. Provision was instead made for the Medical and Dental Council to set fees on condition that it avoided further contracting-out. The threat was that if not successful the Minister would step in to regulate the framework for contracting-out. (Department of Health, 2002).

The Amendment Act, No. 59 of 1984, however eliminated the distinction between contracting-in and contracting-out for doctors and dentists. From that period on any profession or supplier of a service was permitted to set their own fees through their statutory control bodies. The Representative Association of Medical Schemes (RAMS), the statutory association of medical schemes, was also permitted to set a SOB after consultation with supplier representatives. If fees equal to or less than the fees indicated on the SOB were charged the scheme was required to pay the health professional directly. (Department of Health, 2002).

The “Free Market” Reforms: 1984 to 1989

Although medical schemes, due to their close links to employers, were not-for-profit, a commercial arm to the system evolved through the outsourced administration services provided to schemes. As the occupational scheme model makes sense only for large employers, smaller employers found it easier to contract with schemes able to take on more than one employer. Whereas occupational schemes had a natural

sponsor in the form of the employer, schemes serving multiple employers had no natural sponsor other than commercial administrators. Commercial administrators therefore began to set up multi-employer schemes which they could not legally own. The commercial return was derived from the administration fees that were charged to the resulting scheme.

During the 1980s multi-employer schemes became more prevalent. Although de-facto non-profit, their behaviour was commercial in nature. Multi-employer schemes did not however compete for individual members. Instead this was the focus of smaller open schemes and insurance products with the latter explicitly for-profit.

The regulatory framework for schemes was tailored around occupational medical schemes following the governance model of trusts (as with pension funds) and incorporated many protections for members and beneficiaries. Trustees supervised the funds on behalf of specified beneficiaries. As with trusts, schemes were required to apply for registration with a registrar and to file rules that needed to comply with legislated provisions. Key amongst these was a limitation on the right of schemes to vary contributions on the basis of health status or the risk of claiming benefits (or risk rating). Occupational schemes however only varied contributions on the basis of income, with higher income groups required to pay more. In many cases they went so far as to positively discriminate in favour of older members (who on average are higher claimers) who would pay lower contributions than active members.

Schemes were also required to cover certain required benefits, expressed as a percentage of the SOB. In most cases this was 100%.

The emerging commercial multiple-employer schemes however lobbied for more liberalization of the market. They were legally constrained in their commercial activities by two provisions. Firstly, their inability to risk rate created financial uncertainty for schemes unable to predict the risk profile of the employers they were taking on. Secondly, the requirement to offer mandatory minimum benefits eliminated their ability to differentiate benefit offerings for alternative target groups. They consequently began to lobby for legislative change.

The Browne Commission of Inquiry of 1986 largely accepted this commercial case and saw the introduction of risk-rating and cost-sharing with members (co-payments,

through the elimination of mandatory minimum benefits) as a cost saving measure likely to improve the affordability of schemes through greater scheme competition, member behavioural change, and increased scheme consolidation. Medical schemes were consequently deregulated in 1989 with all restrictions on risk rating removed.

During this period deteriorating macroeconomic conditions placed major constraints on public health budgets. This caused noticeable declines in the investment in public hospitals. Although public hospitals technically covered the indigent population only, this was only in relation to financing. Public hospitals as a provider system de-facto covered all income groups. However, budget constraints potentially began to influence public sector career opportunities for specialists, who consequently moved to the private sector, either part or full-time, where they operated as independent practitioners. In many instances these doctors began to set up private hospitals to which they referred patients instead of the public sector. Although this trend pre-dated the 1980s, it was only during this decade that the private hospital system achieved a near complete diversion of private hospital patients away from the state system. (van den Heever, 2012).

Table 1: Private and public hospitals and bed estimates from 1976 to 2010

Year	Private		Public	
	Hospitals	Beds	Hospitals	Beds
1976	25	2 346		
1986	65	6 125		117 842
1989	101	10 936		
1998 ⁴	162	20 908	343	107 634
2010 ⁵	216	31 067	410	88 920

Source: van den Heever, 2012.

The period 1990 to 1994

Two strategic changes to the private health system occurred over the period following the initial deregulation of medical schemes in 1989 to 1994. The first involved the liberalisation of scheme choice for public sector employees, and the second removed mandatory minimum benefit requirements for schemes.

Government medical scheme

The public sector removed the mandatory requirement for membership of designated public sector schemes. This permitted civil servants to choose their own scheme, or

even to choose not to be covered. These formed part of an overall reconfiguration of civil service employee benefits clearly designed to protect civil servants in the post-1994 election period. A fairly substantial market for individual members, targeted by some of the newly emerging open schemes, consequently developed from 1994 onward.

Removal of mandatory benefits from medical schemes

On 1 January 1994 the government removed all requirements for mandatory benefits for medical schemes, consistent with the Browne Commission proposals, and was no doubt hurried through before the change of government in that same year.

The changes impacted during the period from 1994 to 1999, with evident effects on medical schemes, brokers, private providers, and medical product suppliers (medicines, equipment, and consumables) due to the unregulated commercialization of medical schemes and health care providers.

The period of scheme deregulation 1994 to 2000

Medical scheme deregulation, which allowed risk-rating and the removal of mandatory benefits, created a new market for commercial open medical schemes, offering coverage to both multiple employers and individuals. Traditional insurance companies, a commercial group distinct from the traditional medical scheme administrators, saw an opportunity to move into new territory.

Whereas traditional administrators serviced multiple closed schemes, open schemes competed for clients through undercutting community-rated contribution costs by risk-rating, risk-selection, and selective benefit reductions (creating gaps in benefits that largely affected poor risk groups and individuals).

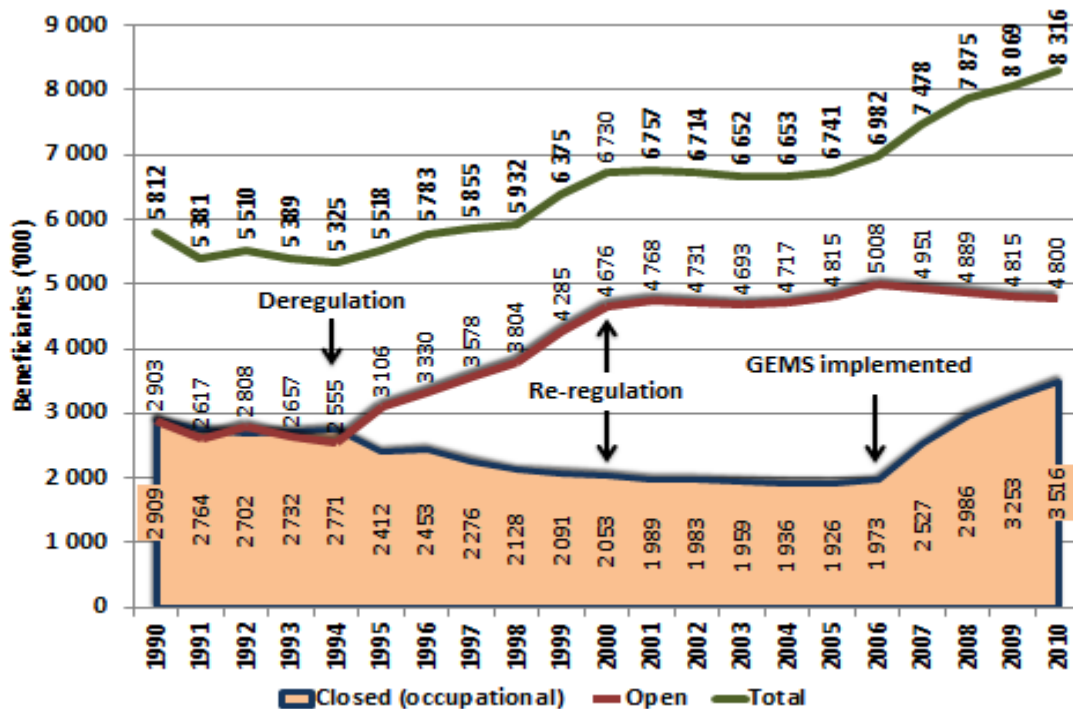
Brokers (financial advisors) serving the traditional insurance and annuity product markets also moved into the market as methods were found to circumvent prohibitions contained in the Medical Schemes Act (of 1967) on their reimbursement.

Payments to brokers occurred either through the co-selling of “health insurance” (insurance regulated through the Financial Services Board) products tailored to coverage gaps in medical schemes (unseen to members); the diversion of administration fees (after at first increasing these beyond the true cost of

administration); the use of reinsurance agreements at inflated premiums to divert a schemes' underwriting surplus to the administrator – which would then be used, in part, to fund brokers; and the charging of contributions at rates different to the registered contribution tables, with the difference, in part, used to fund brokers. (McLeod et al, 2001).

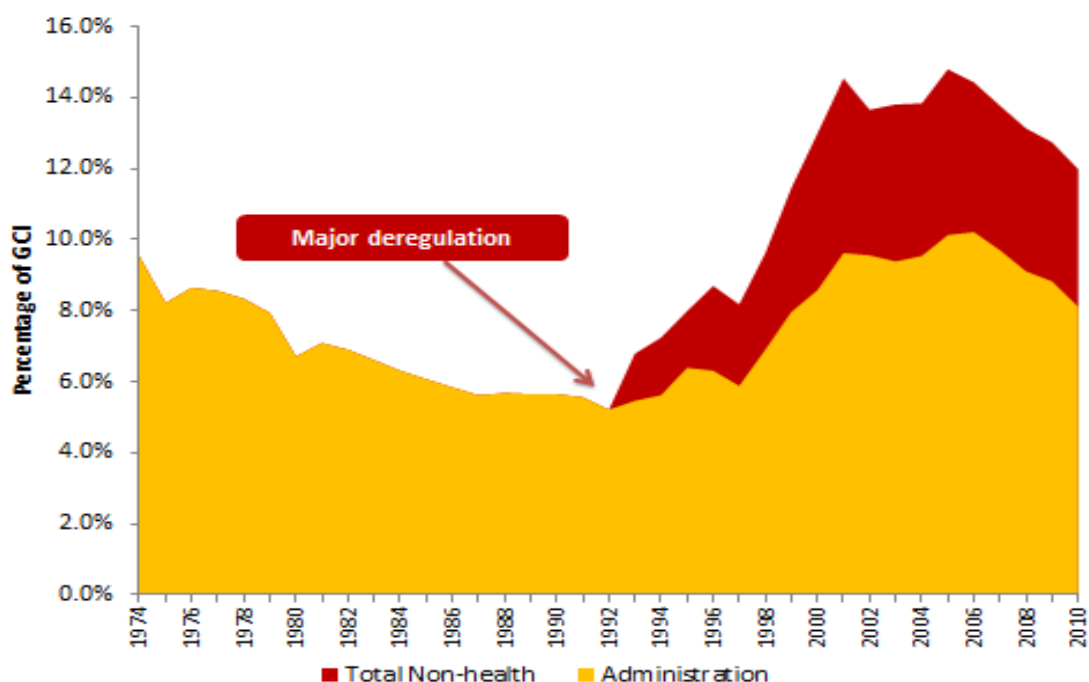
As commercial open schemes were subject to far weaker governance requirements, gaining the approval of conflicted trustees and principal officers was relatively straightforward, all of whom were appointed by the relevant administrator.

Figure 1: Beneficiary changes from 1990 to 2010, distinguishing open from closed (restricted membership) schemes



Source: van den Heever, AM (2012).

Figure 2: Non health expenditure in medical schemes from 1974 to 2010 expressed as a percentage of Gross Contribution Income



Source: Data from 1974 to 1999 compiled by the Centre for Actuarial Research (Department of Health, 2002, p.27); from 2000 to 2010 determined from Council for Medical Schemes (2011) table 24 and figure 35.

The 1994 deregulation unleashed substantial commercial forces which rapidly saw brokers advising employers to close their occupational schemes and join an open (commercial) scheme. A resulting shift of members occurred exactly at the point of deregulation (see **figure 1**) with roughly 2 million beneficiaries transferred from closed to open commercial schemes over the period 1994 to 1999.

This movement reflected implicit competition between administrator types; i.e. those serving the traditional occupational (closed) scheme market and those sponsoring and operating commercial multi-employer or individual (open) schemes. However, business was shifting from the leaner and better priced administrators to far more expensive administrators operating de-facto for-profit schemes. Employers were often persuaded that open schemes were cheaper because they could substantially reduce their liabilities in relation to expensive pensioner groups. From 1994 the introduction

of commercial competing schemes was therefore associated with significant non-health related cost increases (**figure 2**).

Although commercial open schemes were competing on the headline cost of medical scheme coverage, they were never competing on the true cost of coverage. As perceived costs could be reduced by excluding poor risks from cover, open schemes could look cheaper despite higher administration fees, the illegal extraction of profit, and illegal payments to brokers in the form of commissions. The information asymmetry between employers and individual members on the one hand and medical schemes on the other permitted medical schemes to over-price relative to the reduced coverage they were now offering.

Competing schemes probably did not have market power relative to employers and therefore could not increase contribution rates excessively. They could however take advantage of product complexity to extract an excessive margin from an invisible benefit reduction. It is important to note that the benefit reduction implicitly eliminated effective life-cycle risk-pooling (also referred to as cradle-to-grave coverage), which would not be understood by youthful contributors/beneficiaries until they reached the age of retirement (when they would most probably be excluded from coverage or charged a much higher contribution).

The information asymmetry between schemes and consumers also compromised incentives for competing schemes to encourage cost efficiencies on the part of healthcare service providers. As the additional investments to effectively manage health care provider costs would reduce profit margins on administration, schemes preferred simple fee-for-service reimbursement and the associated unsophisticated annual SOB fee negotiations. As a consequence, benefit and contribution costs rose precipitously during the 1990s and into the early 2000s (**figure 3**). This contradicted the expected cost-reduction and efficiency motivations for the 1989 and 1994 deregulation that flowed from the Browne Commission.

Scheme competition on the basis of actual service-provider costs was also risky, as the simple insurance models prevalent within the commercial insurance market (for life, death, and disability) lacked the sophistication to embark on such long-term strategies. The required information systems, provider supervision, and contracting capability involved high up-front investments for an unknown and risky return.

Schemes were also vulnerable to a backlash from providers, particularly specialists which could result in serious reputational damage. Commercial players such as Sanlam and the Southern Healthcare Joint Venture (SHC) with Anglo American failed dramatically during the latter part of the 1990s due to both provider reaction and the failure to properly implement new expensive information systems (see **annexure 1**). Both consequently lost business to schemes continuing to adopt the standard insurance approach and ultimately closed down.

The close cooperation between independent hospitals, which formed the majority of hospitals during most of the 1990s, was coordinated horizontally through the National Hospital Network (NHN) and specialist groups, and was central to the collapse of these attempts to establish selective contracting arrangements. As the hospital market was not concentrated at this point (as this occurred around 2000) foreclosing selective contracts by schemes appears to have required express horizontal collusion between specialists and hospitals (see **annexure 1**).

It can be argued that the express collusion amongst hospitals through NHN during the 1990s has transformed into legal collusion during the post-2000 period through hospital mergers by the three main groups (Netcare, Medi-Clinic, and Life (previously Afrox)). Prior to the market consolidation the main hospital groups (then Medi-Clinic and Afrox) actually saw opportunity from selective contracts, as can be determined from their willingness to enter into these arrangements (**annexure 1**). After 2000, however, the three main hospital groups predominantly impose the same conditions on attempts to establish networks as NHN demonstrates in **annexure 1**, i.e. all or nothing conditions.

“We have succeeded in sticking together and have refused to sign unless this issue of accepting all the hospitals in the NHN have been accepted.”

The “*all or nothing*” requirement however effectively destroys any value that could be achieved from a network arrangement. Networks need to be limited (not wide) to extract a price reduction for the direction of patient volumes to specific facilities. As these contracts would have exposed *inefficient* excess hospital capacity, which relies for survival on schemes’ exposure to unavoidable over-servicing (supplier-induced demand) in the fee-for-service environment, NHN hospitals with their specialist owners would have been forced to rationalize. Importantly, other competing medical schemes would have been forced to adopt the strategies of SHC and Sanlam or lose members to schemes offering better coverage at a lower price. Their failure consequently took the pressure off other medical schemes and providers, defaulting to the preferred, and administratively simple, but structurally inefficient, fee-for-service environment.

Private hospitals appear to have materially improved their market power through consolidation over the period 1994 to 2000 but were still technically not concentrated in the national market, using the Herfindahl-Hirschman Index (HHI) as a measure, in 1999, assuming that the NHN hospitals behaved independently rather than collusively. Nevertheless, over the period 2000 to 2002 the market exceeded a value of 1,800, regarded in the United States as sufficient to trigger an anti-trust investigation (Feldstien, 2005).

Private beds as a percentage of the national total had also increased significantly from 12.4% in 1994 to 17.6% by 2000. Private beds had therefore not only increased in number by an additional 6,659, but this had gone together with a progressive increase in market concentration. A trend that appears to have become systemic, both nationally and regionally, from 2000 (**table 2**). The period of concentration coincides with a structural change in hospital claims costs (**figure 4**).

Table 2: Changes in national hospital market concentration based on acute beds, from 1996 to 2010

Year	HHI	Change	% change
1996	977		
1998	1 161	184	18.8%
2000	1 657	496	42.7%
2002	1 972	315	19.0%
2004	2 139	167	8.5%
2006	2 338	198	9.3%
2010	2 273	-64	-2.7%¹

¹The negative change results from the entry into the market of Hillcrest and Ethekewini hospitals.

Sources: Hospital Association of South Africa, updated with data from the Competition Commission (2012).

The period of medical re-regulation from 2000 to the present

From 2000 onward medical schemes were required to comply with new regulatory provisions and the establishment of a new regulator, the Council for Medical Schemes (CMS). Prior to this period the industry was regulated by a deputy directorate within the Department of Health (DOH) with a staff of seven (Melamet Commission, 1994).

Although the new regulatory framework superficially reintroduced some of the provisions removed in 1994, they took account of the altered landscape. Whereas in 1994 roughly 50% of beneficiaries were in closed (occupational or employer) schemes, now more than 70% of beneficiaries were in open commercially-oriented schemes. Regulation therefore needed to cater for the distinct characteristics of two forms of scheme, one commercial and competing and the other non-commercial, with the former now considerable in size and importance.

The regulatory framework furthermore sought to take advantage of the dynamic characteristics of the market to optimize health-related social protection for income earners. Whereas the Browne Commission recommendations, which were ultimately implemented in their entirety in 1994, gave the market (and more particularly the commercial interests of commercial schemes) greater weight than social protection, the revised framework emphasized both social protection and commercial viability. The evidence furthermore contradicted the Browne Commission presumed market effects of their recommendations. Medical scheme competition, based on risk-rating

and risk-selection, rather than decreasing costs, coincided with structural increases in healthcare costs (**figure 3**).

The new framework incorporated the following strategic elements relevant to market behaviour (Medical Schemes Act, 1998):

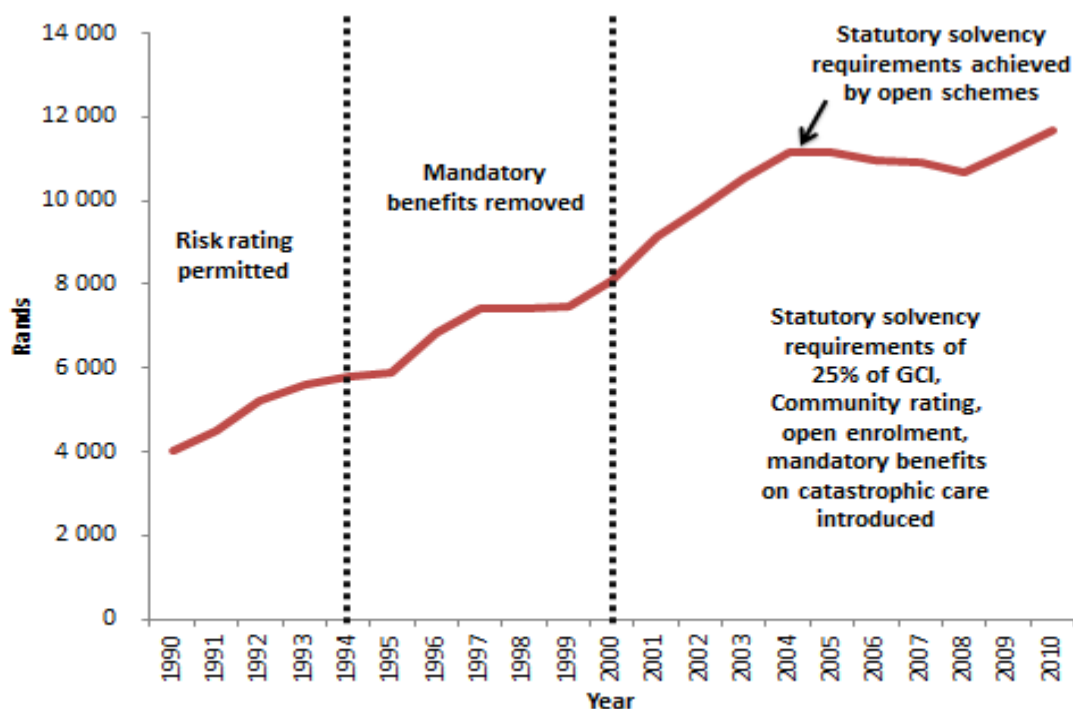
- *Community rating*: which prohibited the setting of contributions on the basis of health status.
- *Mandatory minimum benefits*: which focused on catastrophic health events, including elective, non-elective and chronic care.
- *Open enrolment*: which removed the discretion from any open (commercial scheme) to select which risk groups could join. From 2000 no scheme could therefore risk-select.
- *Mandatory reserve levels of 25%*: which compelled schemes to hold a minimum of 25% of gross contribution income (GCI) as a financial reserve. Schemes were given a five-year period from 2000 to achieve the required reserve. This required that all schemes existing at 1 January 2000 needed to achieve a 25% solvency margin by 2004.
- *Regulation of brokers*: the grey market for advisory services was regularized, but subject to a regulatory framework which required registration, accreditation, and capped fees. Schemes were also compelled to report on broker fees and administrators (from 2004) prohibited from directly or indirectly paying their fees.
- *Governance changes*: required that both board members and principal officers (from 2004) needed to be separated from administrators and/or any entity contracting with the scheme.
- *Establishment of only two forms of scheme*: two forms of scheme were permitted, restricted membership and open schemes. The former required permission from the registrar to exist and could not be constituted merely to exclude poor risk groups or individuals. The latter was the default state for any scheme and no differentiation was made for schemes targeting employers or individuals. This meant that any person could join any open scheme, either through their employer or as an individual.

- *Transparency of costs and scheme operations*: schemes were required to report in detail on their operations, which are provided to the public in the annual reports of the CMS. This allowed all market participants to see the reported costs of competitors.
- *Approval requirements for reinsurance*: which, from 2004, removed the practice of using reinsurance agreements to convert medical schemes into de-facto for profit schemes.
- *Stringent justification for excessive contribution increases*: contribution increases were subject to a stringent annual review by the regulator, with unfair or manipulative changes reduced.

The framework removed a large part (but not all) of the competition based on health status (through risk-rating and risk-selection) and increased competitive pressures to manage costs efficiently. The following effects of the reforms are discernible:

- Real per capita GCI increased on historical trends until 2004, the year when statutory reserves needed (and were largely) to be achieved. Thereafter a structural change occurs to 2010 with contributions largely flattening (**figure 3**).
- Real per capita non-health costs, which were increasing in excess of 20% per annum, flatten from 2001 to 2010 with a real increase over this entire period of only 4.6% (**figure 2**). Beneficiary numbers increase from 6.7 million in 2000 to 8.3 million in 2010, representing an additional 1.6 million lives covered (**figure 1**).

Figure 3: Medical scheme gross contribution income (GCI) per average beneficiary per annum, 2010 prices (Rands) for the period 1990 to 2010



Sources: Council for Medical Scheme annual reports from 1990 to 2011, CPI from Stats SA

- Private health provider and product costs grew dramatically during the 1980s and 1990s, with some stability achieved during the 2000s, but mostly after 2004. Despite similar trends, each cost area follows unique patterns of change. For instance, whereas hospital costs per beneficiary fell below medicines during most of the 1990s, from 2002 they diverge dramatically, reaching twice medicine costs per beneficiary by 2010. (**Figure 4**).
- Specialist costs per beneficiary also fell below medicines from 1981 to 2004, with a 34% differential by 2010. Medicines demonstrate a remarkable real decline in cost per capita, largely attributable to regulatory interventions outlawing inducements paid by manufacturers to doctors, pharmacists and hospitals to drive sales of their products as well as increased co-payments on medicines.

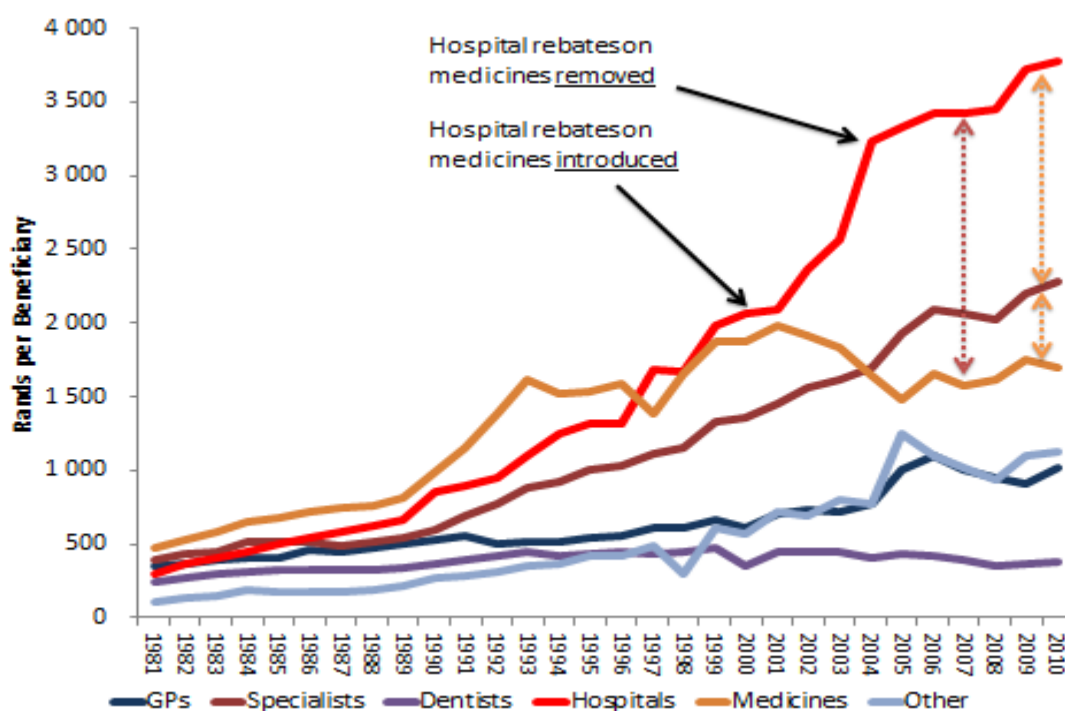
The dramatic divergence between the hospital and medicine per capita costs suggests that hospitals built the profits they were making from the rebates (inducements) into

their facility fees from 2004. The likely market power wielded by private hospitals permitted them to convert one revenue stream into another, with no serious constraint offered by medical schemes. There is no evidence that schemes were able to successfully counter the facility-fee (hospital tariffs) increases.

Objections were raised by schemes who argued that an agreement reached between schemes and private hospitals in 1999 already allowed hospital groups to convert a prior medicine mark-up (introduced during the 1990s) into facility fees (Council for Medical Schemes, 2008). Schemes at that stage wished to remove the perverse incentives associated with the mark-up. However, hospital groups both increased their facility fees (by the value of the mark-up) and covertly introduced the rebate system. This in part explains the dramatic cost increase in hospital costs detectable from 2000. (Council for Medical Schemes, 2008).

The consolidation of private hospitals into major national hospital groups allowed them to implement centralized procurement for medicines, consumables, and equipment. The centralized procurement arrangements, together with significant market power over medical product companies, also facilitated the easy implementation of the rebate or inducement mechanisms (which required product suppliers to increase the prices of their products and rebate agreed amounts to hospital groups to access the now highly consolidated hospital market).

Figure 4: Medical scheme real per capita claims costs from 1981 to 2010 (2010 prices)



Sources: Medical schemes claims data from the Council for Medical Schemes annual reports for 1993 to 2011; GDP and CPI data from Statistics South Africa.

Although medical schemes pay for a wide range of medical benefits, not all are risk-pooled (insured). For insurance to be viable, the insured benefit should: occur infrequently; have a relatively high cost per occurrence; and claims should not be at the discretion of the insured (Barr, 2004). Although these requirements are satisfied for catastrophic health expenses, many healthcare needs occur frequently, involve low costs per event, and are initiated at the discretion of the patient – such as GP care. Medical schemes mitigate this risk by effectively making members self-fund for benefits in this range, with claims exceeding an excess value claimable from a risk-pool.

Benefit expenses recoverable from a risk-pool reflect a very different cost trend to those that are technically funded on a self-insured basis. Hospital and specialist costs are almost entirely funded from risk-pools, while medicines, GPs, and dentists are self-funded through medical savings accounts and annual claims limits (which are

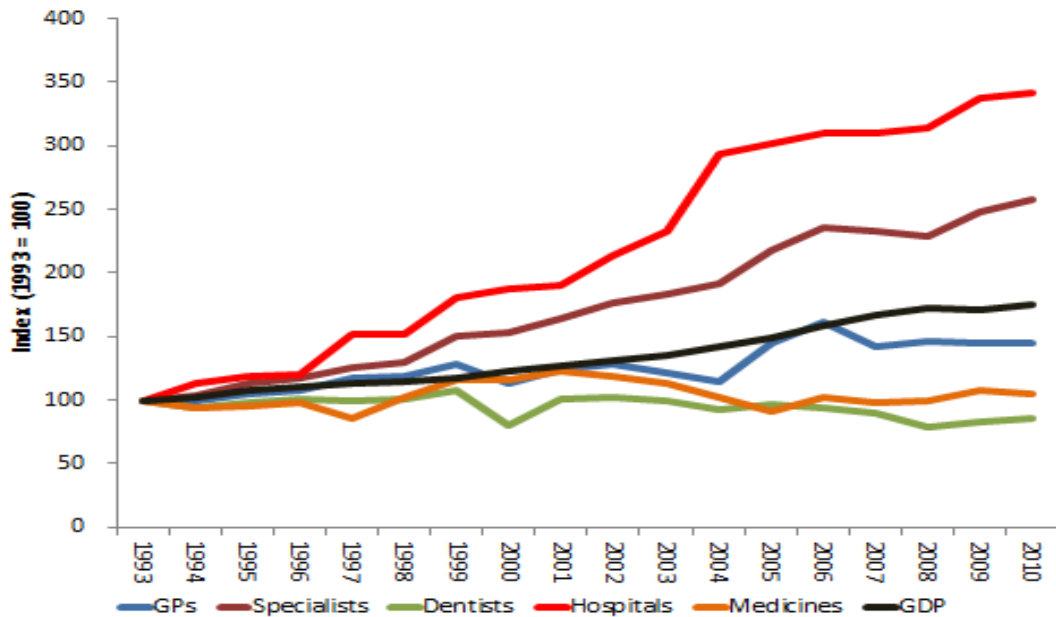
essentially equivalent to an annual budget as scheme members are faced with a use-it or lose-it choice).

Figure 5 illustrates the resulting cost trends, with hospital and specialist costs rising in excess of an index of GDP, with all benefits and GDP at an index value of 100 in 1993, and all other benefits below the GDP index. The difference between hospital and specialist costs is probably attributable to the ability of schemes to shift some part of specialist costs onto medical scheme members by not fully funding their consultation and procedure fees. However, this option does not exist for hospital care, which is typically funded by schemes without limit.

Figure 6 provides the same data as in **figure 5** except that non-hospital benefits are aggregated and medical scheme gross contribution income (GCI) and administrative expenditure (non-health) are added. This shows that from 1993 non-hospital benefits have changed together with GDP, while hospital costs have always grown faster. As the GDP reflects changes in general purchasing power, this suggests that medical schemes have maintained non-hospital costs at GDP growth, but have been unable to contain hospital costs. Schemes have also substituted non-discretionary cover (hospital and specialist) for discretionary cover (out-of-hospital expenses).

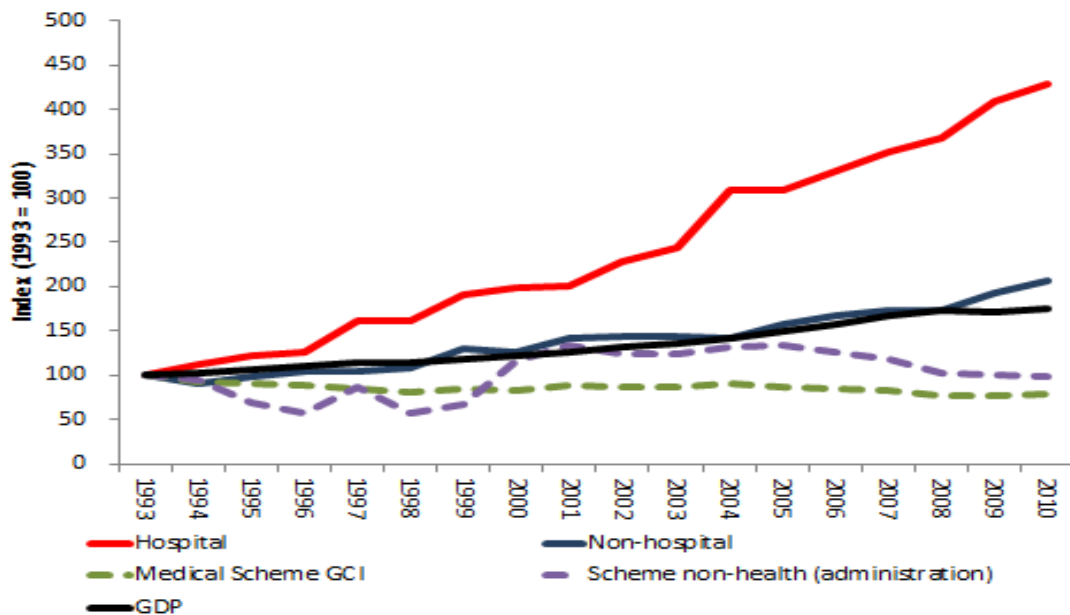
Medical scheme GCI has also grown by less than GDP from 2004 (after the achievement of statutory solvency margins), suggesting that schemes have been under competitive pressure to contain contribution growth despite the hospital and specialist cost increases. This has been achieved by incrementally curtailing non-health costs and discretionary out-of-hospital cover.

Figure 5: Medical scheme real claim cost changes from 1993 to 2010 (pabpa) compared to GDP expressed on a common index with 1993 = 100



Sources: Medical schemes claims data from the Council for Medical Schemes annual reports for 1993 to 2011; GDP and CPI data from Statistics South Africa.

Figure 6: Index changes in real medical schemes expenditure on hospital and non-hospital claims compared to GDP 1994-2010 (1993 = 100)



Sources: Medical schemes claims data from the Council for Medical Schemes annual reports for 1993 to 2011; GDP and CPI data from Statistics South Africa.

A pivotal development occurring during the post-2000 period has been the implementation of the Government Employees Medical Scheme (GEMS). This reversed the 1993 government decision to liberalize the medical scheme participation of public sector employees and their families. Given the scale of the intervention Government chose to implement GEMS on an incremental basis, with all new civil servants incentivized to enrol onto the scheme from 2006, with incentives provided for the rest to either switch from an open scheme or to join. GEMS has grown rapidly from 0 enrolees in 2006 to 1.8 million by 2014 (Council for Medical Schemes, 2014). At its (inevitable) maximum the scheme will exceed 3 million beneficiaries.

GEMS has grown from two sources. New low-income members who were never on a scheme before and existing open medical schemes (those that predominantly target individual members). Large- and medium-sized schemes targeting corporate membership are largely unaffected. However, smaller open schemes have seen their beneficiaries age rapidly with many having to liquidate or amalgamate. GEMS has consequently caused a general consolidation of medical schemes, with only two (Discovery and GEMS) now covering just more than 50% of all beneficiaries. The impact of GEMS on the open schemes' market can be seen in **figure 1**. The increasing concentration of funders could in future reduce competitive incentives to efficiently manage costs and quality.

Conclusions

The system of private financing has changed considerably over the more than 100 years of its development. Influences have included changes in: the public health system; demographics; the macroeconomy; technology – including funder administration, health services and health products; and regulation. Key changes in the environment allude to probable causal relationships with market outcomes, in the form of costs and quality.

Various regulatory changes have altered the factors over which competition occurs, changing funder incentives to either disengage from, or confront the performance of the market for healthcare goods and services. If funder incentives to manage costs and quality are weak, providers of healthcare goods and services will be free to conduct their market activity inefficiently.

However, even if funders are adequately incentivised, market power imbalances between funders and producers of healthcare goods and services could avoid competitive pressures by successfully influencing the contracts they enter into by limiting them to those that favour their interests. As funders have not been effectively incentivised to purchase healthcare goods and services efficiently, even since 2000, it is not possible to conclude whether poorly incentivised contracting or excessive provider market power best explains the cost trends over the past 25 years. It is however very plausible that both factors play a considerable role.

Explaining hospital and specialist cost increases over the past two decades

Various factors appear to have contributed to private healthcare cost increases over the past 25 years. Per capita expenditure has increased for two reasons: firstly, prices have risen structurally over time, with real increases invariably in excess of inflation (Council for Medical Schemes, 2008); and secondly, the volume of consumption (demand), in the form of additional activities has risen structurally, particularly hospital-based activities (Council for Medical Schemes, 2008). The simple question is whether consumers would have made these purchases if they had another choice.

In markets for other product types, consumers can choose between substitutes or defer consumption of a product when its price increases, or where the total expense to a household is high. In healthcare this option is removed for insured purchases. Decisions are made for the consumer, by conflicted market participants, and demand is substantially accommodated by a third-party. Under such circumstances it is plausible to conclude that new investments that expand the productive capacity of private health system will face lower levels of market-related risk (for over-capitalisation) when compared to a competitive market.

A reduced risk of investment should influence the price at which investments are purchased and their volume. The cost of investment is likely to be structurally higher, as the user of the investment, for instance the hospital group, can adjust both prices and demand to achieve a targeted return on investment, regardless of the inefficiencies in investment and subsequent delivery. Changes in demand would therefore be expected to follow changes in the supply of services. Whereas a conventional monopolist's market power would be reflected in its super-normal

profits, and be based on charging a higher market price, here increased profits can arise from continuous annual increments in demand based on low risk investments in new supply, purchased at inflated market prices. Profit margins alone, which increased structurally over the past 15 years (Felet et al, 2012), would not fully reflect the likely profit maximisation strategy.

Industry-sourced groups argue that the private health market is efficient and competitive (Netcare, 2014). Changes in costs and activities are argued to derive from: aging medical scheme beneficiaries; technology change; legislative requirements placed on schemes to cover certain conditions; nursing cost increases resulting from public sector increases; and HIV and AIDS (Council for Medical Schemes, 2008). However, as the medical schemes population has virtually not aged from 2000 to 2012 (from 2000 to 2012 the average beneficiary age increased from 31.2 to 32.0) (based on Council for Medical Schemes Annual Reports from 2000 to 2012) it is implausible that utilisation changes have arisen from morbidity changes due to aging. HIV and AIDS, which is at very low prevalence levels for medical scheme beneficiaries (Council for Medical Schemes, 2008), will not have structurally increased hospital admissions, as successful out-of-hospital treatments keep hospital admissions low. There is furthermore no evidence that nurse cost increases have any impact on final hospital costs (Council for Medical Schemes, 2008).

For supply-induced demand to become a driver of cost, consumer choices would need to be largely captured by product suppliers. The information asymmetries open the door for product suppliers, of both health insurance and healthcare goods and services, to take control of the consumer decisions via conflicted agents (brokers and doctors). Three important choices are removed for consumers to act on: price differentials (between healthcare goods and services and health insurance); cost differentials (between healthcare goods and services); and quality differences (healthcare goods and services and health insurance).

Medical doctors, who are responsible for most of the critical resource utilisation decisions in the private market, presently take no cognisance of price, cost and quality information. As such, neither do the consumers. Even the choice of doctor is devoid of any relevant choice criteria.

Within an insured market, doctors have no incentive to compete based on price or quality (both of which are in any case explicitly prohibited by the Health Professions Act) for two reasons: consumers have no incentive to choose doctors based on price as they are insured against the expense; and secondly they cannot make choices on quality due to the absence of relevant information. Doctors have also successfully resisted market interventions that attempted to introduce price and quality competition through concerted conduct (see **annexure 1**).

Failure in the market for hospital services is also indicated by what it competes on. Instead of competitively bidding for market share from medical schemes, they instead compete to attract specialists with established practices. As specialists drive demand from the supply-side, any need to compete on cost or quality for market share is circumvented.

Attracting specialists involves offering inducements, practice support and expensive equipment. Whereas bidding for market share from schemes would at the very least result in competition on costs, competing for specialists requires cost increases through (riskless) over-capitalisation and the over-supply of specialists. Although it is argued that there is an under-supply of specialists in South Africa (Netcare, 2014), the opposite is quite plausible. Evidence of over-capitalisation is indicated by the fact that the South African private sector has more beds per population than the UK and US (**table 3**), and more magnetic resonance imaging and cat scanners per million people than most industrialised countries (Council for Medical Schemes, 2008).

Table 3: Medical scheme beneficiaries and private hospital bed to population ratios for 2001 to 2010

Year	Medical scheme beneficiaries ¹	Bed to population ratios (per 1,000) ²
2001	6 971 627	3.43
2002	6 992 407	3.44
2003	6 855 471	3.55
2004	6 898 107	3.86
2005	6 741 231	4.03
2006	6 981 724	3.93
2007	7 442 515	3.87
2008	7 709 697	3.76
2009	7 946 218	3.78
2010	8 243 527	3.77

The US and UK acute bed to population ratios in 2010 stood at 2.9 according to the OECD³

Sources: ¹Council for Medical Scheme Annual Reports for the periods 2001 to 2010; ²Hospital Association of South Africa Annals for various years; ³OECD health data.

At the level of insurance, consumers are entirely dependent upon a system of conflicted advice for the medical scheme and other health insurance products offered. These products are not standardised and cannot be compared by consumers without expert advice. Benefit levels can be subtly varied between seemingly similar products through strategic exclusions, while prices (premiums/contributions) vary for demographic reasons, rather than the cost and quality of healthcare goods and services. Brokers are paid by the insurance product suppliers, and are therefore incentivised to make sales for the insurers who fund their businesses.

Medical practitioners (GPs and specialists) are also deeply conflicted. They have shares in private hospitals, directly and indirectly receive inducements from private hospitals and medical product suppliers – in particular medicines. Hospitals, who are also able to influence the sale of medical products, have in the past received substantial inducements from medical product suppliers (it is not clear whether this practice has been discontinued).

Inducements flow to agents within the health system able to influence the final demand for products. They would plainly have no economic rationale directed at actors unable to directly or indirectly influence demand. The fact that private hospital

groups attracted billions of Rands worth of inducements until 2007 (Council for Medical Schemes, 2008), in the form of rebates, is indicative of their influence over the final demand for these products (as doctors would only have the products available that the hospitals chose to supply).

Inducements, by serving as methods used by product suppliers to capture demand, implicitly also serve as proxy indicators of those actors that are able to materially influence demand. Completing this thought process, inducements could also be paid by product suppliers to final consumers to attract demand – in the form of price discounts or competitive quality improvements. The limited role played by final consumers in making product decisions indicates why no such offers are made in practice. Final consumers are largely irrelevant, and consequently ignored. Whereas legitimate inducements to final consumers take the form of competitive prices, inducements to proxy consumers (the agents) take the form of price surcharges – increasing prices and costs.

That no behavioural consequences flow from the prevalent degree of consumer ignorance and manipulation in the South African private health market is inconceivable. The continuous rise in costs is therefore most probably explained by the market imperfections that result in the near total capture of the consumer. The cost increases incorporate both price and volume (demand) changes, with all forms of efficiency compromised. Furthermore, quality of care, worryingly, is invisible, and therefore not a basis for competition. To date no hospital or set of providers has seen any advantage in unilaterally competing on quality, suggesting an implicit consensus amongst providers not to rock the boat.

The theory of market failure in context

Medical schemes

Medical schemes have weak incentives to compete on valid market criteria that would be most relevant to consumers. Presently, no contracts of insurance between medical schemes (and any other insurer for that matter) and consumers fully internalises the costs and the quality of health goods and services.

While it could be argued that there is some partial internalisation of cost, largely attributable to Medical Schemes Act (1998) which, to a degree, prohibits competition

on health status, there is no internalisation of quality. As a consequence, no contracts between medical schemes (or any other insurer) and providers of health goods and services, fully internalise cost and quality.

Weak medical scheme contracts with providers of health goods and services allow for excessive costs to be passed on to medical schemes – which schemes accommodate financially through a combination of price increases (contributions/premiums) and benefit reductions. As consumers principally seek medical scheme coverage for catastrophic health expenses (mainly hospital and specialist services), coverage of these benefits is protected relative to other benefits. This results in an upward sloping demand curve for hospital and specialist services, and a relatively inelastic, but downward sloping demand curve for other benefits.

Within the “other benefits” are GP contracts, which, rather than being used to strategically filter purchases of specialist and hospital services, are retained in simple fee-for-service arrangements. Little attempt is made to require that GPs manage the catastrophic care decisions, or to contractually remove damaging conflicts of interest. This purchase is simply one of those that are incrementally downgraded by schemes to protect the coverage of unmanaged catastrophic specialist and hospital expenses.

Historical trends suggest an inelastic and slightly upward sloping market demand curve for medical scheme cover. The ease with which medical schemes, at a market level, retain demand, regardless of real contribution increases, would support the argument that they have weak incentives to aggressively internalise cost and quality in their contracts. It is easier and safer to make incremental adjustments to invisible (to consumers) benefits rather than antagonising comfortable relationships with providers of health goods and services.

Hospitals and specialists

For hospital and specialist goods and services, the price, demand (volume) and quality of healthcare goods and services are determined independently of each other at the discretion of providers of healthcare goods and services. Demand can also be expanded at the discretion of the supplier, independent of market prices and consumer household budget constraints.

A change in demand is therefore not a function of changes in market prices, nor is a change in prices driven by changes in demand. Both price and demand are set to achieve remuneration targets regardless of their costs of production. In the case of specialists, these can be set through collusive arrangements, or independently. Hospitals are able to make these decisions strategically, at a corporate level, due to their market power.

The incentives of hospital group owners are to maximise profits through a combination of price increases and supply-induced demand (in excess of what a competitive market would permit). They are able to achieve this through exercising market power in relation to funders, who have weak incentives to purchase healthcare goods and services efficiently, easily restricting their contracts to arrangements which externalise cost and quality of care risks back to final consumers.

General practitioners

General practitioners technically hold a strategic position in the health market, as they can heavily influence specialist and hospital referrals. However, presently they have no incentive to operate as strategic gatekeepers and agents for consumers, as this depends heavily on funders entering into sophisticated contracts with them to fulfil this role. Medical schemes however have weak incentives to deepen their contracting with GPs (discussed above).

Other healthcare products

The survival of healthcare product lines supplied into the private healthcare system depends not on the choices of final consumers, but on proxy consumers (doctors and hospitals). A sophisticated system of inducements is used to incentivise the proxy consumers to drive demand, regardless of the price (but not of the inducement).

Final product prices are consequently a mark-up on the prices that would be expected in a competitive market – i.e. where the final consumer receives the inducement. Product quality is also invisible to the final consumer, and market choices are therefore likely to be perverted. Although certain of these products may fall foul of the incremental benefit reduction process, to the extent that they form part of captured purchases, such as doctor's prescriptions and medical tests, the market affect is similar to that of an insured purchase (particularly as many insured consumers only

find out about the co-payment or balance-billing requirements after they've received treatment).

Resolving market failure in the South African private health system

Market failures in the South African private health market appear structural in nature and are therefore unlikely to be addressed through conduct-related interventions (i.e. through the Competition Act). The failures derive principally from the structural exclusion of the consumer from all strategic choices. All negative cost and quality changes in the system stem from this central cause.

For the market to be corrected, structural measures, sufficient to re-position the location of the consumer such that all market decisions flow from their informed decisions, are required. A focused range of regulatory interventions could be implemented that selectively remove certain factors from competition and replace them with factors over which competition must occur. These need to be applied across the entire market to address the market failure.

As it is unlikely that the information asymmetry that exists between patients and suppliers of healthcare goods and services can be addressed, a properly incentivised agent is essential. As healthcare purchases are inevitably bundled, and access to care is made affordable only through insurance, the health insurer is a pivotal consumer agent, provided their incentives are aligned. Aligning incentives requires that schemes compete on a deeper contract that is transparent to consumers and derives from their revealed preferences.

The contract with consumers needs to include the benefits package, the costs of healthcare goods and services and the quality of healthcare goods and services. This can be achieved through a requirement for all schemes to offer a standardised minimum benefit package including all essential coverage, and a non-standardised benefit package which can contain discretionary cover. The standardised mandatory benefit package would be the same in all schemes, allowing for ease of comparison by consumers.

As package prices vary according to demographics (more or fewer old beneficiaries), resulting price (contribution/premium) variations can be eliminated through prospective risk-equalisation on the standardised minimum benefit package. Risk equalisation is a regulatory intervention that involves the formula-based transfer of funds from schemes with excess good risks to those with poor risks such that all schemes face the industry risk profile. Such formulae are at an advanced stage of development in South Africa – and are administratively straightforward to implement.

Whereas standardised and risk equalised minimum benefits would materially enhance price competition between schemes, it would not resolve concerns about the quality of care purchased. Consistent with the findings on the NHS reforms (Gaynor, 2012), for quality to be internalised as a factor over which competition occurs, information on quality needs to be given to all strategic decision-makers in the system. This would require that quality outcome information (for instance, hospital-based mortality rates, outbreak information, and specialist practice outcome indicators) be reported publicly, and available to all the strategic agents and final consumers.

Strategic agents would include all brokers, medical schemes, GPs and final consumers. As it is likely that final consumers will not process such indicators routinely, agents acting on their behalf need to be incentivised to provide truthful advice on these indicators. This can only occur if all conflicts of interest have been removed. Furthermore, it would be essential that provider agents (GPs) cannot be bypassed as advisors, which ensures that quality, as a factor driving competition, results in behavioural changes amongst relevant market actors.

Where benefits have been standardised, schemes will have to differentiate themselves on two basic criteria: the price (contribution/premium), which will be heavily influenced by contracts they set up with healthcare providers; and the quality of care of the service providers they contract with. If brokers are required to be independent, they can play a key role advising on the quality of the provider networks contracted with for the standardised package. Members would not need input from an agent to understand price differences, as these will be transparent. They would however need to understand whether there is any trade-off between price and quality.

Introducing the above reforms would substantially influence the incentives of medical schemes to compete aggressively on cost and quality. This would be achieved by removing demographics and package variations from competition, establishing a market for un-conflicted independent advice on scheme choice, and making provider quality of care information public.

Correcting scheme incentives to contract efficiently is however unlikely to prove sufficient to correct the market as a whole without addressing the strategic role of doctors, and in particular GPs. General practitioners need to become the strategic filters for referral to higher levels of care. For this to occur they firstly need to be independent (without any conflicts of interest) of all other providers of health goods and services; and secondly, be required by law (for a scheme to be permitted to reimburse) to be the first point of call in the health system.

This, combined with the public availability of strategic information on referral service outcomes, would establish the foundation for efficient contracts of any type between schemes and providers. Schemes would be in a position to establish networks based either on explicit cost and quality, or via GP contracts where they filter referrals based on publicly available information and any additional criteria established by the scheme.

Explicit prohibitions on arrangements that establish conflicts of interest between specialists and other healthcare goods and services would also separate out the markets for specialists from that for other services. This combined with clear prohibitions on collusive conduct, should prevent the foreclosure of pro-competitive contracts emerging between schemes and specialists. As medical practitioners have great difficulty working out when they've crossed a competition boundary – a framework of guidelines for practitioners should be developed and enforced by the competition authorities.

To the extent that efficient hospital contracts are prevented from entering the market through the exercise of market power by hospital groups, two strategies are implied: divestiture; and/or the establishment and application of a well-designed licensing system. Measures falling within the latter area should be considered before the former, as both structure and conduct issues can be addressed simultaneously.

The licensing criteria can regulate the: the allowable contracting modalities; restrictions on anti-competitive conduct; and the manner in which they engage with GPs and specialists. Where conduct falls short of the license requirements, it can be suspended, withdrawn and/or a penalty imposed. Divestiture on general grounds is potentially less useful than the threat of divestiture based on conduct transgressions linked to well-developed licensing criteria.

As it is inevitable that fee-for-service contracts, which include nothing but price in the agreement, will form an important part of both the insured and out-of-pocket market, market power imbalances in the setting of these prices also need to be addressed. As these prices cannot reasonably be set through negotiations with individual consumers or schemes, a transparent and representative central bargaining arrangement is required to set final prices for hospitals and medical professionals. These tariffs should apply as a default where no bilateral contract is negotiated between any healthcare service provider and a scheme.

Concluding remarks

This paper has attempted to construct a theory of market failure in the South African private hospital system. Although far from complete, this analysis has attempted to demonstrate that market-based solutions to healthcare problems are conceivable, provided the problem is treated as an engineering problem rather than a crude choice between state and market.

A key challenge is whether government actors have the capabilities required to address a market problem holistically, which requires multi-faceted regulatory reforms implemented with a concerted purpose. If government understands this potential, it opens the door to highly leveraged reforms which can rely on markets to achieve important social ends.

Failing which, government can default to the heavy-handed levers, vulnerable to multiple governance risks and failures, and the contingent risk that both government and market mechanisms structurally fail indefinitely into the future.

References

Barr, N. *Economics of the welfare state*. Fourth edition. Oxford University Press. 2004.

Brown Commission. *Final Report of the Commission of Inquiry into Health Services*, 1986, RP 67/1986.

Felet, A. Kishman, D. Fiandero, F. *Do hospital mergers lead to healthy profits?* Paper produced for Wits Competition Conference, Unpublished. 2012.

Groenwald, B. Stiglitz, J.E. “Externalities in economies with imperfect information and incomplete markets. *Quarterly Journal of Economics*. 1(2 May), 1986, pp.229-64.

Rothschild, M. Stiglitz, J.E. “Equilibrium in comprehensive insurance markets: an essay on the economics of imperfect information.” *Quarterly Journal of Economics*. 90, 1976, pp.630-49.

Cooper, Z. Gibbons, S. Jones, S. McGuire, A. *Does Hospital Competition Improve Efficiency? An Analysis of the Recent Market-Based Reforms to the English NHS*: Centre for Economic Performance, London School of Economics and Social Science, CEP Discussion Paper No 988, June 2010.

Council for Medical Schemes. *Evaluation of Medical Schemes’ Cost Increases: Findings and Recommendations*: Policy Brief 1, March 2008.

Council for Medical Schemes. *Annual Reports for the periods 1991 to 2013*.

Department of Health. *Inquiry into various Social Security Aspects of the South African Health System*: based on the health subcommittee findings of the Committee of Inquiry into a Comprehensive System of Social Security, 2002.

Department of Justice, Federal Trade Commission. *Statements of Antitrust Enforcement Policy*: 1996.

Dranove, D. Shanley, M. “Price and Concentration in Hospital Markets: The Switch from Patient-Driven to Payer-Driven Competition”: *The Journal of Law & Economics*, vol. XXXVI, April 1993, pp179-204.

Dranove, D. White, W.D. "Price and Concentration in Hospital Markets, The Switch from Patient-driven to Payer-Driven Competition": in *The Journal of Law and Economics*, vol. XXXVI, April 1993, pp.179-204.

Dulleck, U. Kerschbamer, R. "On Doctors, Mechanics, and Computer Specialists", The Economics of Credence Goods: *Journal of Economic Literature*, Vol. XLIV, March 2006, pp. 5-42.

Feldstien, P.J. "The Market for Hospital Services": Chapter 11 in *Health Care Economics*, 6th edition, Thomas Delmar learning, 2005, pp.258-295.

Krishnan, R. "Market Restructuring and Pricing in the Hospital Industry": *Journal of Health Economics*, vol.20, 2001, pp.213-237.

McLeod, H.D. Slattery, P.G. van den Heever, A.M. "The Use and Abuse of Reinsurance in Medical Schemes". Presented at the Actuarial Society of South Africa annual convention, October 2000. Published in the *South African Actuarial Journal*. Vol. 1, 2001.

Melamet Commission. *Commission of Inquiry into the manner of providing for medical expenses*. 1994.

Pauly, M. "The economics of moral hazard: comment." *American Economic Review*. 58(3, part 1), 1968, pp.531-37.

National Treasury. *Budget Review 2012*. 2012.

Netcare. *Overview paper – submission to the health market inquiry*. Non-confidential version. 2014.

Organisation for Economic Co-operation and Development. Working Party No.2 on Competition and Regulation:

- Cooper Z. *Competition in Hospital Services*: 13 February 2012, 2012b.
- Gaynor, M. *Competition in Hospital Services*; 13 February 2012, 2012c.

Melnick, G.A. Zwanziger, J. Bamezai, A. Pattison, R. "The effects of Market Structure and Bargaining on Hospital Prices": *Journal of Health Economics*, vol.11, 1992, pp.217-233.

Office of Fair Trading. *Report on the market study and proposed decision to make a market investigation reference*: December 2011.

Propper, C. Burgess, S. Green, K. "Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market": *Journal of Public Economics*, Vol.88, 2004, pp.1247-1272.

van den Heever, AM. "The role of insurance in the achievement of universal coverage within a developing country context, South Africa as a case study." *BMC Public Health*, Suppl 1, 2012, pp.1-13.

Vogt, W.B. Town, R. *How has hospital consolidation affected the price and quality of hospital care?* Robert Wood Johnson Foundation. The Synthesis Project. 2006.

Annexure 1: Private doctor and hospital response to schemes' attempts to introduce selective contracting in 1997

Private hospitals fight for survival

Just how fast the health market moves was illustrated last month with the negotiations between the National Hospital Network (NHN) and Southern HealthCare (SHC). The one day the NHN, which consists of 96 hospitals and clinics owned mainly by doctors, was locked in battle with SHC. The parties professed not to speak to each other.

The NHN made a complaint to the Competitions Board against SHC for refusing to allow their members to utilise hospitals of their choice, and doctors who worked in these hospitals refused to sign provider contracts unless the managed care company also included NHN in its hospital network. Then first 14 East Rand doctors resigned from the provider network, followed by another hundred doctors or so.

At that stage SHC chief executive officer, Kathy Walstead-Plumb, said she wanted to contract only with some of the individual NHN hospitals in some geographical areas. The NHN's "all-or-nothing-attitude" when it comes to signing up hospitals in their network, does not help, she said.

Two days later the whole scenery changed. The NHN won the battle. The SHC relented under pressure from the doctors and all the NHN hospitals were put onto its network.

But behind this lies the fact that the fight for survival of private hospitals and clinics is on and hospitals not included in the networks of big managed care companies will soon start feeling the pinch. At the moment about 65% of the 20 000 private hospital beds are filled, and the expected shift from public-sector hospitals to the private sector is not apparent.

Both Sanlam Health and SHC have contracted with both Apex (Afroc and Medi-Clinic) and Clinic Holdings on a per diem rate. About 70% of hospitals in these two groups are in urban areas and this is where the battle for survival will take place as the shift of patient volumes

between hospitals will follow the contracts.

SHC initially chose only those hospitals that suited its needs, while Sanlam Health decided in January that its strength lies in casting its net as wide as possible. At that stage Sanlam Health clinched a deal with the NHN to provide services for Sanmed in certain areas, while providing services for Transmed and Bestmed in the rest of the country, ensuring the most extensive hospital network in the fledging local managed health care industry.

By contracting with NHN the company secured the goodwill of doctors working or holding shares in those hospitals, which was a good strategic move. Another move was to agree to develop hospital quality criteria for accreditation, tender procedures and patient preferences together with the NHN. "This will help us to better understand what is important in a hospital-doctor and hospital-patient relationship", said Dr Louis Rossouw, Medical Director, Networks at Sanlam Health.

A month later, SHC followed suit.

What seems to be the lesson from this? Peter Botha believes doctors have done very well by refusing to allow the big managed care organisations to divide them. "We have succeeded in sticking together and have refused to sign unless this issue of accepting all the hospitals in the NHN has been resolved".

For the moment everything seems rosy coloured again, but until more patient volumes are brought into the private sector from the public sector the future of many private hospitals is bleak.

Source: "Private Hospitals Fight for Survival", Managed Care Review, published in the South African Medical Journal, March 1997.