



**NEW ZEALAND INSTITUTE FOR THE STUDY
OF COMPETITION AND REGULATION INC.**

**Paying for the Hospital Waiting List Cull at the GP's
Surgery:
The Changing Locus of Financial Risk-Bearing in New
Zealand's Primary Health Care Sector**

May 2007

Bronwyn Howell

New Zealand Institute for the Study of Competition and Regulation Inc. and
Victoria Management School, Victoria University of Wellington, PO Box 600,
Wellington, New Zealand. Email bronwyn.howell@vuw.ac.nz

***Acknowledgement:** The author wishes to acknowledge the helpful comments provided by Lewis Evans, the support of Glenn Boyle and the New Zealand Institute for the Study of Competition and Regulation, whilst undertaking research into New Zealand primary health care markets and the Deane Endowment Trust for financial support during the preparation of this paper. Any errors or omissions remain the responsibility of the author.*

Abstract

This paper uses an analysis of the treatment of ‘random’ and ‘predictable’ risk in primary health care capitation contracts to critique the arrangements in the New Zealand Primary Health Care Strategy (NZPHCS) introduced in 2002. The New Zealand capitation contract arrangements are unusual, in that the capitation funder does not meet the full costs of the commissioned care, the independent, private sector practitioners receiving capitation payments are able to charge the patient for any costs not met by the funder, and the capitation incentive becomes increasingly sharp over time as the funder selectively prioritises different population groups to receive greater proportions of its subsidies.

The NZPHCS arrangements appear to have assumed away the reallocation of random risk by treating the capitation subsidy as a ‘notional averaged treatment subsidy’. Whilst increased government funding has reduced out-of-pocket costs for patients, total system costs have undoubtedly risen relative to a fee-for-service counterfactual, with the burden of extra costs falling disproportionately on the most frequently ill. The vulnerable individuals for whom the system was supposed to provide greater benefits are also consequently disproportionately affected by random shocks that affect the demand for general practitioner services. The costs of shocks such as the waiting list cull are certainly being paid at the GP’s surgery, with the sickest paying most of these costs, and the habitually well reaping the benefits.

Introduction

“There are many mechanisms for paying physicians; some are good and some are bad. The three worst are fee-for-service, capitation and salary”. (Robinson, 2001:149)

Recent years have seen the emergence of many health care policies employing risk-sharing tools aimed at aligning the activities of service providers with the objectives of purchasers and funders. An example is managed care, where responsibility for managing all aspects of the financing and delivery of health care for a defined set of patients is assigned to specific care management entities (Robinson, 2004; Rivers and Tsai, 2003). Whilst managed care models involve the use of many non-price elements (e.g. utilisation review), a common characteristic is some element of financial risk-sharing (‘supply-side cost-sharing’ – Ellis and McGuire, 1986).

The agency theory rationale for supply-side cost-sharing is that by contractually linking the care manager’s or care provider’s financial returns to a proxy for the achievement of some desired characteristics, or by decoupling financial returns from undesirable characteristics, the desired activities will be pursued and the undesirable ones eschewed. Effort exerted by the manager or practitioner, which is unobserved and unobservable by the policy-maker or insurance purchaser setting the contract terms, will be redirected towards the pursuit of desirable, rather than undesirable, activities (Holmstrom and Milgrom, 1991; Robinson, 2001). Commonly-used supply side cost-sharing contracts are price-and-volume, partial and full capitation (Danzon, 1997).

Partial or full capitation contracts have become common in primary health care remuneration as they reduce the over-consumption externality of subsidised health care by financially penalising practitioners who deliver ‘too many’ consultations. Capitation is the principal form of government-funded remuneration for England’s Primary Care Trusts (Keen, Light and Mays, 1999), is common in insurance-funded remuneration of primary care physicians in the United States (Robinson, 2004; Hagen, 1999) and is the basis of New Zealand’s partially tax-funded primary health care strategy (King, 2001). They are also extensively used in public sector ‘outcomes-based’ (Honore, *et. al*, 2004) and ‘performance-based’ (Martin, 2002) contracting. Reduced emphasis on the consultation as the primary payment determinant is attributed with a shift in primary care delivery focus away from interventions in the event of illness towards the promotion and maintenance of wellness (Coster and Gribben, 1999; Cumming, 1999; Malcolm, 1997). They are also attributed with being stimuli for increased equity, targeting of high health need, encouraging a team approach to primary

health care, and a change in focus towards a care management model as opposed to a model of episodic intervention with a focus on illness (Crampton, Sutton and Foley, 2001).

However, capitation contracts have limitations. Robinson (2001:4) identifies that capitation is inferior to fee-for-service in that it does not recognise the extent of practitioner effort exerted: “its payment is determined prospectively without regard to the number of services provided, overpaying physicians who stint on care and underpaying those who provide many complex services”. Capitation also performs poorly in regard to risk management, as it is “imperfectly adjusted for the severity of illness of each covered patient. Even a well-adjusted capitation payment rate fails to compensate physicians who treat patients whose condition deteriorates, leading to greater utilization and cost, for reasons independent of the physician’s own actions” (*ibid*). Howell (2007) suggests that a high dependence of practitioners’ incomes on unpredictable variations in individuals’ demand for services poses particular challenges for the use of the instrument in health care markets generally, and primary health care markets in particular, that are not present in many other markets.

In section one, this paper explores the effect that random events can have upon practitioners’ incomes in capitation contracts generally. Section two then discusses the particular arrangements of the New Zealand Primary Health Care Strategy (NZPHCS) capitation contracts introduced in 2002. The effects of the New Zealand strategy are illustrated in section three with two recent ‘random’ events: a strike by junior doctors in the country’s hospitals, and a decision by District Health Boards (DHBs) to remove all individuals on waiting lists for secondary and tertiary (hospital-provided) consultations and procedures for six months or longer back to their primary care providers for ongoing management. The section also illustrates how other elements of the New Zealand capitation contract, such as the requirement that increases in government capitation subsidies be passed on to patients in the specific capitation classes as reduced patient payments, and rigorous enforcement of practitioner charges to patients may affect sector outcomes. Section four concludes.

1. Capitation in Primary Health Care Markets

In a partial capitation contract, the remunerated party’s income comprises a fixed component f determined by the ex-ante characteristics of the population for whom care responsibility is assumed, and a variable component v for each unit of output q produced at a cost c (profit = $f - q(c-v)$). In a full capitation contract ($v = 0$), the recipient’s income is invariant to the number of units of output produced. By decoupling remuneration from cost drivers, the higher is f and the lower is v , and the greater the recipient’s reliance upon income from capitation than upon

the number of units of output produced. The recipient now faces incentives to reduce costs in order to increase profitability – for example, producing fewer outputs or finding cheaper ways of producing them. Behavioural change ensues because the undesirable behaviour is no longer rewarded, or because the rewards will accrue only by engaging in the desired behaviours (Milgrom and Roberts, 1992).

However, capitation contracts have complex effects on practitioner behaviour, as they share two specific types of financial risk between the purchaser and the provider: ‘predictable risk’ and ‘random risk’. ‘Predictable risk’ is the responsibility for managing those identified cost-causing behaviours that are the specific target of the capitation contract, such as a provider propensity to order too many treatments, more expensive pharmaceuticals or unnecessary laboratory tests (Zeckhauser, 1970), or the use of preventative treatments to improve patient health state thereby reducing the number of consultations required (Crampton, Sutton and Foley, 2001). ‘Random risk’ relates to those factors for which the service provider assumes either partial or complete financial responsibility via the risk-sharing contract, but is powerless to control, such as events that cause an uncontrollable increase in demand for consultations (e.g. a localised epidemic) and discrepancies between the ex ante assessment of demand for services by the individuals for whom responsibility to deliver care is assumed, and their actual demand (‘random’ demand variation).

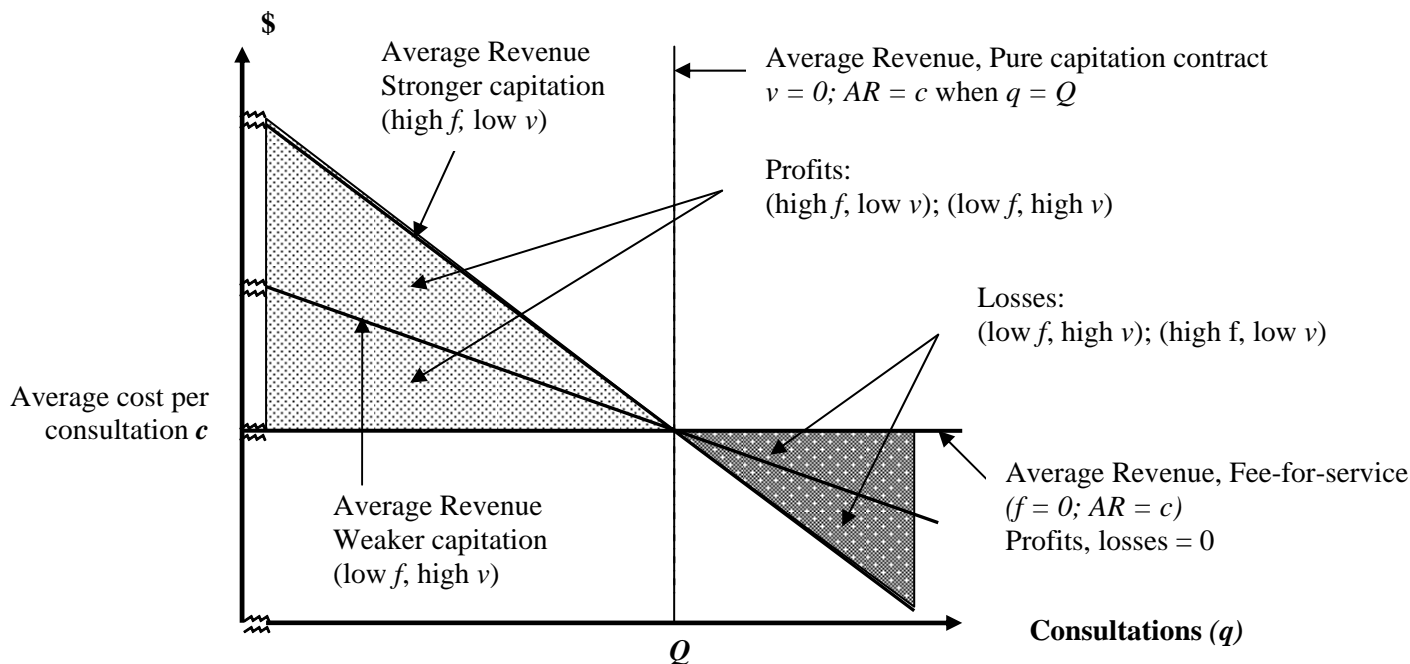
When a capitation contract shares responsibility for ‘predictable’ risk management between the purchaser and the provider, it also shares an equal amount of ‘random’ risk, which also affects provider income. The effects may be either positive or negative, but importantly, they are outside the control of the practitioner. It is therefore crucially important in a capitation contract that the amount of financial risk shared is carefully balanced. The stronger the capitation contract incentive (that is, the higher is f and the lower is v), the greater the proportion of random risk that is shared, and the more the practitioner’s income comes to depend upon random factors that cannot be controlled. If the ‘random’ effects on the practitioner’s income are large compared to the ability to manage income via the ‘controllable’ factors related to ‘predictable’ risk, then irrespective of the amount of effort exerted pursuing the desired activities, the practitioner’s income becomes essentially a ‘lottery’. The incentive to pursue the desirable behaviour is ‘crowded out’ and so the desired activities are not pursued. Rather, the practitioner will exert effort instead in activities that maintain or increase income given the amount of random risk assumed (e.g. ‘cream-skimming’ to ensure that the patients for whom care management responsibility is assumed are more likely to bring higher fixed income and lower demand for services).

Howell (2007) posits that health care capitation contracts pose a particular difficulty in that the parties sharing both types of risk- third party insurers and governments – derive their justification for undertaking the contracting responsibility in the first place from the fact that they are more efficient managers of individuals’ uncertainty in the costs arising from uncertain demand for demand for health care services than the individuals themselves (as per Arrow, 1963). When third-party risk aggregators share the responsibility for ‘random’ risk that they are charged with managing more efficiently to smaller provider risk pools via capitation contracts, the benefits from reduced costs of ‘predictable’ risk must be traded off against the increased costs of ‘random’ risk being less efficiently managed than within a single large pool.

Moreover, even more caution must be exerted in the case of primary care capitation contracts, as the repeated interaction between specific individuals and practitioners leads to serial correlation between unknown cost-causing patient characteristics not captured in the compensation agreement (e.g. unknown genetic predisposition in an individual) and practitioner income. The consequence is the emergence of habitually profitable and habitually loss-making practices (‘lucky’ and ‘unlucky’ respectively). The greater the amount of risk shared (i.e. the higher is f and the lower is v), the greater the difference in profitability between the practices of different types, and the greater the propensity for undesirable behaviour to be pursued at the expense of desirable behaviour, in order to maintain long-term practice financial viability or to maximise profits.

Figure 1 illustrates. Assume it costs an average practice c to deliver an average primary care consultation (including all overheads and a fair return on the human capital and time invested by the practitioner). Average revenue received per consultation is $f/q+(v-c)$. Under a fee-for-service contract charged at cost ($f=0$; $v=c$), the practitioner makes no profits and no losses on all consultations delivered (i.e. ‘breaks even’). The number of consultations delivered is independent of q , and is determined solely by the practitioner’s willingness to work. Under a capitation contract, however, the practitioner receives $v<c$ for each consultation delivered. There will be a maximum number of consultations that the practitioner will deliver (Q), at which point the practice ‘breaks even’. If ‘luck’ (i.e. demand arising from the actual health state of the population for who capitation is received) means that the number actually delivered q is less than Q , the practice makes a ‘windfall’ profit. If luck means $q>Q$, the practice makes real financial losses.

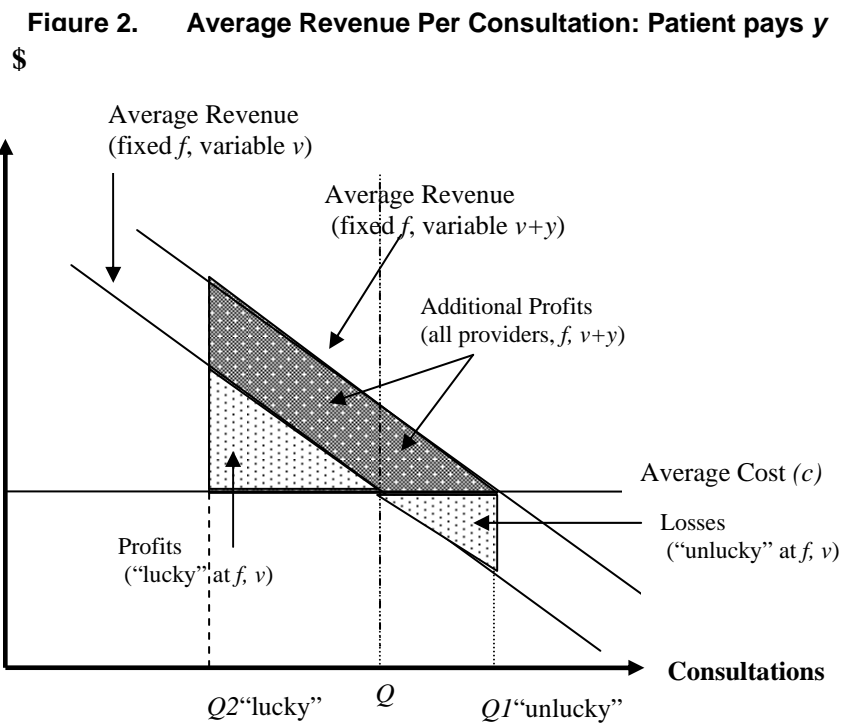
Figure 1. Average Revenue Per Consultation: Standard Capitation



A capitation contract incentivising desirable behaviours looks for providers delivering more than Q consultations through their own choices to reduce the number to Q in order to remain financially viable, or for those with costs higher than c to reduce them to c . However, the same contract will financially penalise those practitioners with costs c and not over-producing, whose ‘luck’, rather than behavioural choices, leads to more than Q consultations being delivered. If these practices are to remain financially viable, they must reduce costs below c (e.g. shorter consultations), ration services (e.g. institute waiting lists) or pass the extra costs on in some other way (e.g. institute a patient payment y in addition to v). The sharper the capitation contract (i.e. the lower is v and the higher is f), the steeper is the slope of the average revenue curve (Figure 1), the greater the profits and losses, and the greater the additional costs that must be borne by the patients of ‘unlucky’ practices ($q > Q$) relative to the patients of ‘lucky’ practices ($q < Q$). Ironically, these practices are ‘unlucky’ in the first place because they have a patient base with higher demand (i.e. ‘sicker than average’), so luck combined with the capitation contract leads to sicker-than-average patients bearing more of the consequences of ‘luck’, in either lower care quality (shorter consultations, waiting lists) or higher prices (y charged to them) than the ‘healthier-than-average patients of ‘lucky’ practices. Howell (2007) notes that budget financing (e.g. as used by government departments to finance fully-owned care providers) is simply the extreme case of full capitation, where $v=0$.

Figure 2 illustrates the effect of being able to charge a patient y in order for ‘unlucky’ providers to recover the costs of ‘bad luck’. Assume that the ‘unluckiest’ practitioner not

over-producing at cost c must deliver QI consultations to meet demand at f and v . This results in the practice making losses of y per consultation. If the practice charges patients y per consultation to break even, the practice will still produce QI consultations (technically, the practice's average revenue curve moves upward) as illustrated in Figure 2). However, if one practice can charge y , all other 'luckier' practices will also charge y , meaning that the average revenue curve of all practices moves upward. 'Unlucky' practices delivering between Q and QI consultations without charging y now make profits instead of losses, and the 'lucky' practices producing fewer than Q consultations make higher profits than before. The total number of consultations produced is higher than anticipated by the capitation contract which was used to set f and v . However, quality is not compromised. The cost of the additional quality provided (equal to the reduction in quality faced by the patients of 'unlucky' practices if y could not be charged) is the area represented by the additional profits plus the losses of the 'unlucky' provider in Figure 2.



In effect, the ability for a capitated provider to charge patients for the costs not met by the capitation contract effectively 'undoes' the motivation for using financial incentives to alter practitioner behaviour in the first place. If the practitioner can shift the costs of 'luck' onto the patient, the costs of 'predictable' risk can also be shifted. Financial incentives are typically used where the desired behaviour is costly to observe, or direct observation is infeasible (as in third-party purchasing). In such circumstances, the desired changes are difficult to induce using non-financial instruments (Milgrom and Roberts, 1992). Allowing

practitioner charging in a capitated contracting environment thus achieves none of the desired behavioural changes, and results simply in the total costs of the same number of consultations provided being higher than in the case of fee-for-service contracting.

2. The New Zealand Primary Health Care Strategy

The NZPHCS was introduced in July 2002 (King, 2001). The key financial instrument is the introduction of capitation payments to Primary Health Organisations (PHOs), who then contract with primary care providers to deliver services to patients. Practically all primary care providers are paid via a ‘back-to-back’ contract that sees the capitated ‘GMS/Nurse’ payment being ‘passed through’ directly to care providers (Howell, 2005). A key motivating factor for the strategy is to increase the share of government funding in the sector from between 30% (Austin, 2004) and 40% (King, 2001) in 2001 to around 80% as occurs in the secondary and tertiary sectors (Howell, 2007a).

2.1 Service Providers Set and Charge v (plus y)

The first defining feature of the NZPHCS is that, although the government share of primary care sector costs is increasing, patients are still required to pay the balance of costs not covered by the government payments to service providers. That is, whilst the government sets and pays f , the service providers set v and the patients pay it. This arrangement has prevailed in New Zealand since 1938, when an agreement between the government and service providers allowed service providers to charge patients directly for any costs not covered by government subsidies (Ashton, 1999). This charging arrangement has been ‘grandfathered’ through into the NZPHCS (Howell, 2005), despite the fact that the government payment has changed from an ex post treatment subsidy paid to targeted individuals (albeit collected by their service providers on their behalf to defray costs) to an ex ante capitation payment made to service providers in recognition of the anticipated (not actual) demand of the patients for whom responsibility has been assumed (Howell, 2007a).

In the three years following the introduction of the policy, government funding for primary care increased by 43% (Hefford, Crampton and Foley, 2005). However, the government’s total share of expenditure did not increase by as much as the Minister had expected given the substantial additional injection of government funding (King, 2004). From Figure 2, this is most likely occurring because loss-making providers have been able to ‘pass on’ the costs of risk management to patients. Whereas the government payments are set using a ‘notional averaged variable payment’ v for an average number of consultations Q , most ‘unlucky’

practices have been able to set the patient payment to $y+v$, passing on the costs of bad luck to patients, and avoiding severe financial distress.

Evidence that such practices are occurring and are likely widespread is implied in the observation by CBG (2004) that very, early into the operation of the system, providers not yet receiving capitation payments for patients were charging less on average for consultations for identical patient groups than those providers receiving capitation subsidies. CBG (2004) finds this observation puzzling, but interpreted as per Figure 2, such an observation would be expected, as uncapitated practices do not face the costs of ‘random’ effects, as they continue to be remunerated under fee-for-service contracts. Furthermore, McAvoy and Coster (2005:11) report that “many GPs have already benefited, reporting improved financial incomes and financial benefits from joining PHOs”, suggesting both that profits are being accrued by ‘lucky’ providers and even ‘unlucky’ providers are avoiding severe financial stress by shifting the additional costs of financial risk bearing (both ‘random’ and predictable) onto patients.

2.2 Increasing Incentive Strength Over Time

The second defining feature of the NZPHCS is that the government capitation payments are increased over time according to a set of political priorities determined by government budgets. The first beneficiaries were individuals aged over 65 and under 5, and those patients registered in PHOs where more than 50% of the registered individuals were in the lowest income deciles or of Maori or Polynesian ethnicity (“Access Practices” – the remainder are termed “Interim Practices”). Over subsequent years, higher subsidies were extended to 5-14 year olds, 15-24 year olds and 45-64 year olds. The final group, 26-44 year olds, will receive higher subsidies from 1 July 2007, by which stage it is anticipated that most differences in the funding of Access and Interim practices will be eliminated.

Each capitation increase has also been associated with political expectations that any increase in government payments will be directly matched by corresponding reductions in the fees charged by practitioners for the subsidy class receiving the more generous government assistance. For example, the Prime Minister stated that the July 1 2006 capitation increases meant that “700,000 people aged between 45 and 64 would now pay \$27 less for doctor visits and \$3 for prescriptions – instead of \$15 – if they enrolled with a primary health organisation”¹. Despite more groups becoming higher-capitated, there is still an expectation

¹ <http://www.stuff.co.nz/stuff/0,2106,3719646a7144,00.html>

that some groups will continue to pay different patient payments based on class membership – for example, there is strong pressure for practitioners not to charge a fee for consultations delivered to children under 5 years old. This group commands the highest capitation payments, but generates the greatest demand. DHB and MoH policy material supports the contention that this approach to fee setting is expected for price monitoring and if necessary, price regulation (DHB, 2006).

The overall effect of this requirement is that, over time, the average amount of a practitioner's income that comes from government subsidies (f) increases and the share from patient payments (v) decreases. The capitation incentive strength thus becomes increasingly sharp each time the government includes another group of the population amongst the higher-capitated (as per Figure 1). At each increase, the share of both 'predictable' and 'random' financial risk that practitioners are required to bear becomes greater. United States evidence suggests that the amount of primary health care demand variation that can be reliably be predicted using individual characteristics such as age, gender, past health consumption, income and ethnicity is between 20% and 25% (Newhouse, 1006; Robinson, 2004). Ma and Riordan (2002) report that significant practitioner behavioural changes have been evidenced in the case of contracts with very weak capitation incentives. Yet, the New Zealand contracts will ultimately see practitioners facing extremely strong incentives, where 80% of their income is fixed, and only 20% comes from patients. That is, the income earned from delivering an additional consultation will be only 20% of the cost of delivering it.

2.3 Implications

The NZPHCS is paying practitioners using a capitation instrument, but is requiring practitioners to charge different levels of v for each subsidy group depending upon the level of the capitation subsidy paid. Whereas on the one hand, capitation contracts are paying 'insurance premium subsidies' to insurer-managers to manage individual demand risk, on the other hand, practitioners are required to charge patients as if the government payments are 'notional averaged treatment subsidies' as occurred under the previous fee-for-service system. These arrangements take no cognisance of who will bear the additional risk management costs imposed by the system design.

The peculiar charging arrangements and political expectations limit the ability for practitioners to 'average out' the differences in demand facing the practice by using 'gains' from lower-than-expected demand from one subsidy group to cross-subsidise higher-than-expected demand from another subsidy group. The result is that the effects of 'random risk'

are being shared not simply with 4000 practitioner pools of 1200 to 2000 patients, but with up to 48 different patient pools within each practice pool (Howell, 2005). Any of the benefits of aggregated management of individual demand uncertainty via a centralised risk management (insurance) pool has effectively been dismantled.

As illustrated in Figures 1 and 2, and discussed above, the random risk component is almost certainly being passed onto patients in the form of higher patient payments, with much of the additional government subsidies being extracted by providers in the form of profits largely resulting from 'luck'. However, the allocation of the additional costs of the system is quite inequitable, as it results in those patients potentially most in need of care facing a greater share of the additional out-of-pocket or care quality reduction costs. Moreover, the additional risk management costs are borne in proportion to the health state of the individual. Someone who is frequently ill must make more visits, so pays a greater number of payments towards the additional risk management costs than a less frequently ill individual. Indeed, those who make no doctors visits pay none of the additional risk management costs, despite receiving the benefit of a reduced expected cost of treatment if, in the future, care is required.

The overall effect is a primary health care system that is likely very much more costly than the pre-2002 counterfactual, with a very inequitable allocation of the additional costs, due to the dismantling of any centralised patient demand risk management instruments. It will also be unable to utilise financial incentives to achieve any of the practitioner behavioural changes articulated in the policy, such as an increased emphasis on preventative care, as the ability to pass through costs to patients removes any ability to control practitioner behaviour by constraining practitioner incomes. Whilst price controls may limit the ability to pass through risk management costs, such actions will simply accelerate the effects of inequitable allocation of resources by forcing a reduction in the number of consultations provided, and the conversion of the additional financial costs currently being accrued into reduced care quality for those most likely to need multiple consultations.

3. *Illustration by Stylised Examples*

This section uses a few stylised examples based on actual events to illustrate the effects of the allocation of random risk under the New Zealand arrangements as discussed in Sections 1 and 2. The first examples take a truly 'random' shock to demand facing a primary care practice – a strike by hospital junior doctors and the decision by District Health Boards to 'cull' hospital waiting lists by requiring all individuals unable to be seen by a hospital provider within six months to be removed from waiting lists and referred back to general practitioners for

management. These examples show that ‘luck’ is not just confined to the initial allocation of patients amongst practices, but that the New Zealand arrangements also spread the cost of other ‘bad luck’ unequally amongst practices and patients. The second set of examples illustrates the perverse allocative effects of the translation of capitation payments by government into ‘notional averaged treatment subsidies’ for the purpose of setting and regulating patient charges. Such an arrangement ignores how a practitioner will allocate the costs of risks that are unable to be controlled by clinical practice amongst the balance of the patient base. These examples illustrate how a change in capitation payments will necessarily be associated with cost increases for those patient groups not receiving capitation increases, reflecting how the risks are now shared amongst very small patient pools (the practice) rather than larger pools as would occur with a centralised risk manager.

3.1 Doctors’ Strike/Waiting List Cull

The recent junior hospital doctors’ strike and implementation the policy to remove all patients unable to be seen by hospital specialists in less than six months from hospital waiting lists represent ‘random’ events not foreseen by either party when the original capitation contract was agreed. In both of these instances, the patients whose demand for hospital services was unable to be met in the tertiary system were referred back to their primary care providers for treatment. These policies thus comprise a ‘shock’ to demand for primary care services. Each patient referred back by the DHB would require at least one, but presumably several, additional primary care consultations that the practitioner would be unable to avoid providing, relative to the number that would have been provided had the patients been treated at the hospital.

Such ‘uncontrollable’ increases in demand impose the full costs of the additional consultations upon providers. Under fee-for-service, practitioners would be indifferent to such an event or policy change. Under capitation, however, a practice incurs the full costs of an additional consultation, but receives no additional government payments. Moreover, as the capitated patient is charged a fee less than cost ($v < c$), the additional income will be insufficient to cover the additional costs of the unanticipated consultations. The practice now incurs a deficit in respect of the additional demand. The deficit can be covered in one of three ways: by passing on the costs to patients via higher patient charges, reducing total costs by shortening all consultations, or by ‘cream-skimming’ – for example, refusing to treat DHB-referred patients or declining to register patients who are likely to need hospital treatment in the future, and might be subject to delays in receiving this treatment due to the vicissitudes of the hospital systems.

3.1.1 The Base Case

Assuming that consultation time reduction and patient cream-skimming are initially eschewed, the only option left is to raise fees (*patient payment* = $v+y$). A simple numerical example illustrates. Suppose each consultation costs the practice \$50, and the patient payment (v) is initially set at \$20. Each additional patient referred back by the DHB imposes a \$30 deficit on the practice. The more patients referred back by the DHB, the greater the number of consultations, and the greater the deficit incurred. The practice can break even only by charging the additional costs to all patients via an increased patient payment y . All patients now pay the risk premium. Practices with more, and sicker (i.e. requiring multiple consultations), patients referred back must charge their patients more (i.e. more payments of y each are required) to recoup costs than those with fewer patients referred. If higher patient charges cannot be levied, then the costs are incurred as reductions in service quality, with the patients of the 'unluckiest' practices (i.e. those with more patients or higher-need patients referred back) receiving the greatest reduction.

Thus, all sick patients will pay more at the GP's surgery to meet the costs of risk and demand shifting imposed by the DHB referral policy, but the patients of GPs with more and sicker patients referred back will pay proportionately more of the costs of the DHB policy. Note that well individuals pay none of the additional costs of the policy, as they make no GP visits. This is in direct contrast to the pre-2002 system, when all individuals, sick or well, shared the additional costs of a reallocation of the demand of a subsidised individual between the tertiary and primary sectors, via taxes covering the increased number of treatment subsidies paid.

3.1.2 Different Capitation Classes and Patient Charges

Now consider the effects of the mandatory requirement that individuals for whom higher capitation payments are received when well must also pay commensurately lower patient payments when, they fall ill and seek a primary care consultation. Assume the practice in the base case charges high-capitated over 65 year-old patients \$15 per consultation, and low-capitated 25-44 year-old patients \$48 per consultation. The deficit for each additional DHB-referred over-65 consultation is \$35 and for each additional 25-44 consultation \$2. The higher the proportion of higher-capitated (ex ante assessed as 'high needs' or 'high political priority') patients referred back under the DHB policy, the greater the deficits incurred by practices and the higher the commensurate price increases to all patients must be. The patients of practices with large numbers of higher-capitated patients referred back will pay proportionately more of the cost of the DHB policy than those of the practices with fewer higher-capitated patients referred back.

The costs of the DHB waiting list cull, in particular, are unlikely to be trivial. Typically, higher-capitated elderly and young people are disproportionately represented in hospital waiting lists. Thus, the probability of a waiting list patient referred being a high-capitation individual will be substantially higher than the probability of the patient being a low-capitation individual. These individuals are also likely to require multiple additional primary care consultations in order to continually treat the waiting list condition and to reassess the eligibility for re-entry into the waiting list system. These non-trivial additional costs, imposed by the DHBs shifting demand and financial risk into the primary sector, will have a non-trivial effect upon the costs paid by all sick individuals seeking primary care.

This example illustrates the problem of treating the capitation subsidy as if it is a 'notional averaged treatment subsidy' when setting v for each subsidy group. Government capitation payments are paid ex ante in respect of all individuals, irrespective of their actual demand for care. Treatment subsidies are paid ex post in respect of individuals who have fallen ill and need treatment. Whilst a capitation subsidy can be mathematically averaged out to create a 'notional average treatment subsidy', this will match a patient's actual demand only if the individual consumes exactly the amount of care expected in the average payments. Each patient receiving treatment may have an identical need for treatment, but the government payment is confusing a risk-rated insurance premium subsidy, as occurs with a capitation payment ex ante, with a socially-motivated wealth transfer when requiring an equally ill individual of a different subsidy class to pay a different patient payment ex post, as was the case pre-2002 when the government payment was a targeted treatment subsidy transferring wealth from the whole population to targeted individuals.

In the NZPHCS, risk management costs become bundled in with wealth transfers when the number of consultations consumed by each subsidy class differs, for genuine, random risk-related reasons, from the anticipated number used to set the 'notional averaged treatment subsidy'. The result is that individuals will pay different amounts based upon practice risk characteristics, and that the direction of the payment size may be opposite to that intended by the wealth transfer. The 'solution' clearly is to separate risk management from wealth transfers by making the payments for patients receiving identical care (e.g. zero, as in England's NHS, or a standard fee for each patient receiving identical Accident Compensation Corporation -funded treatments from the same providers) with the distributional effects of wealth transfers being accounted for at the point of subsidy (i.e. all individuals receive a capitation premium subsidy, so all individuals should contribute a top-up ex ante to ensure that the risk management costs are shared equally).

For example, extending the ACC analogy, a forester and an office manager both pay the same patient payment at the GP for treatment for an identical injury, even though the forester's employer pays a higher (risk-rated) premium subsidy ex ante. The forester and the clerk both 'top up' the costs of the ACC liability via income-related premium top-ups paid as payroll taxes. Wealth transfers (from high-income office managers to low-income foresters) and risk management (risk-rated premiums reflecting higher anticipated injury rates for foresters) both occur at the point of the subsidy payment, meaning that further wealth transfers do not need to be made at the point of treatment (although if the ACC payment is less than the provider's costs, they may both have to make an identical payment to the provider to meet the total fee charged). Thus, health/injury state does not get confused with either risk management or wealth transfers when ex post when each individual has an equal claim on the available treatment resources (i.e. injured clerks do not have to pay more simply because more than the average number of foresters is injured, whilst uninjured clerks and foresters pay nothing at all, as is the (translated) case under the NZPHCS).

3.1.3 Patient List Mix

Assume now that one practice has a high proportion of individuals 'on the books' assessed as eligible for more generous capitation subsidies because they exhibit the ex ante characteristics deemed more likely to result in higher consultation demand (Practice A), and another practice (Practice I) has a low proportion. Practice A thus receives a higher proportion of its revenue from fixed payments, and it can only charge its patients on average low values of v . Practice I receives a lower proportion of its income from fixed payments, and more from patient payments (i.e. v is higher).

Both practices provide the same number of consultations (K) in a given period. The average patient payment v at Practice A is \$20 per consultation, and at Practice I \$35. Assume also that, as a consequence of the DHB patient referral policy, each is required to provide an additional 200 consultations in a given period. Practice A incurs a deficit of $200 \times \$30 = \6000 , and Practice I a deficit of $200 \times \$15 = \3000 . If each practice spreads the additional costs across all patients, Practice A patients will pay an additional $\$6000/(K+200)$, twice the increase faced by Practice I patients, $\$3000/(K+200)$. Thus, the DHB referral policy results in the ex-ante assessed 'higher need' Practice A patients, who are deemed less able to meet the costs of higher patient payments, and more likely to be dissuaded from seeking primary care by the size of the co-payment, and ostensibly more likely to need care in the first place, facing higher patient payment increases than the 'lower priority' Practice I patients.

This example illustrates numerically the greater risks to financial viability posed by sharper incentives. Practice A faces a sharper incentive contract than Practice I. The financial costs of a random shock are much greater under sharper incentives. It also illustrates the wealth transfer inequities resulting from treating ex ante premium subsidies with ex post treatment benefits when setting patient payments. These are not equivalent as a consequence of random risk elements. This example shows the inequitable effects if the ex ante premium subsidy is actually a fair reflection of actual demand.

3.1.4 Regulatory Intervention in Practice Price Setting

It now becomes clear how critically important the ways in which primary care patient charges will be regulated under the New Zealand capitated system become. Unless regulators are fully informed about all of the components leading to changes in patient out-of-pocket payments levied in all capitation classes by all providers, naïve price controls threatening the financial viability of practices may be imposed, accelerating the development of a two-tier system differentiated on patient risk, care quality and ownership.

Under the capitated system, if price changes are regulated as if they are still solely performing the role of topping up a ‘notional average treatment benefit’ per consultation without taking account of changes in financial risk for ‘random’ effects, then regulators may unwittingly prevent the fair recouping by practitioners of legitimate additional costs of the capitated system in order to maintain financial viability. DHB materials and politicians’ expressed intentions suggest that neither group fully appreciates how the capitated system has shifted substantial amounts of additional random risk costs onto practices relative to the pre-2002 system. There is consequently substantial additional regulatory risk in the New Zealand system that ‘naïve’ price regulation might impact adversely upon the ongoing financial viability of general practices. Thus, it is not surprising that GPs have expressed significant concerns about the prospect of being subject to price regulation by DHBs under the currently-voiced understandings².

Assume now that a naïve regulator, knowing that Practice A receives higher capitation payments, and therefore, in line with treatment benefit pass-through expectations, must charge lower patient out-of-pocket payments than Practice I, is faced with the price rises in example 3.1.3 above. If the regulator presumes that the capitation payment is simply a ‘notional averaged treatment subsidy’, then the price increase of $\$6000/(K + 200)$ imposed by

² See, for example, GP Leaders’ Forum. 2006. Auckland GP meeting told to reject govt plan. <http://scoop.co.nz/stories/GE0605/S00162.htm>

Practice A is 'unreasonable' given that Practice I with the same demand increase imposes a price increase of only $\$3000/(K + 200)$. If the regulator restricts the price increase by Practice A to that imposed by Practice I, then Practice A becomes financially unviable.

Clearly, in a capitated system it is quite inappropriate to treat capitation subsidies as 'notional averaged treatment subsidies' for the purposes of either setting patient payments or the pursuit of equity based on actual health state.

3.2 Selective Capitation Payment Increases

The potential for naïve imposition of price regulations in the absence of detailed understanding about the allocation of 'random' and 'predictable' risk is further illustrated by the consequences of requiring increases in capitation payments paid in respect of well individuals (f) to be passed through directly as commensurate decreases in patient payments for individuals seeking treatment (v). Assume that a practice has 1000 patients. For simplicity, assume that initially, none are subsidised via capitation; however the same effect will apply to all patients not included in each capitation increase, including those already highly-capitated. The practice charges \$50 per consultation, which equals costs (including a fair return on investment). On average, each patient makes 4 visits a year (4000 consultations). Practice costs are $\$50 \times 4000 = \$200,000$.

Assume now that half the patients (Group Y) are provided a capitation 'subsidy' of \$100 per year. The remainder (Group X) are funded as before. The capitation funder requires that the practice reduce the patient charges for Group Y to \$25 per consultation. On past consumption, this would appear, on average, to be fiscally neutral for the practice. However, as Group Y patients now pay less than the full cost of a GP treatment, they increase their consumption to five visits per year as a consequence of paying lower prices. Whilst some of the increased demand may be for justifiable medical conditions, some will be increased moral hazard over-consumption by the 'worried well'.

As a consequence of the increased consumption, the practice costs now increase to $\$40 \times 4500 = \$225,000$ per annum. Group Y patient payments yield $\$25 \times 2500 = \$62,500$, which added to the capitation income contributes \$112,500 to the practice's income. As Group Y charges are capped at \$25, the practice must recoup all other costs from Group X patients. The Group X patient payment is now $\$112,500/2000 = \56.25 per consultation. The capitation pass-through requirement for Group Y patients means that the patient payment for

Group X patients, who receive no subsidy increase, will rise in order to pay for the increased consumption of the newly-subsidised.

This example further illustrates how sensitive any price regulation in the capitated system must be to the trade-offs between random and predictable risk. Even though allowances may be made in the price regulations for some extent of increased consumption it is extremely difficult for a regulator to ascertain the different likelihoods of patients in different practices, or even within different capitation categories in the same practice, to alter their demands in response to marginal changes in patient payments. Two practices with identical patients in all other respects may levy different price changes in response to changes in capitation payments because one has more 'worried well' on the books than the other. As price regulation will prevent any shifting of these patient-induced costs on in patient charges, they become a cost to the practice and a risk to its ongoing financial viability. Selectively increasing patient subsidies increases practice costs, so must be considered when setting incentive strengths if the outcome is going to raise efficiency, and when regulating prices in order to maintain practice financial viability.

Given that over half the population received no treatment subsidies pre-2002, there must be considerable uncertainty about their propensity to over-consume when provided with what appears to them in the perspective of their experience as notional treatment subsidies (reduced out-of-pocket payments) under the capitated system. International evidence suggests that over-consumption by the worried well is substantially more likely to occur amongst higher-income and higher-educated individuals than amongst lower-income, less well-educated ones (Shen and McFeeters, 2005). As income was the primary determinant of subsidy provision pre-2002, it is likely that the progressive inclusion of high-income individuals in the capitated system will result in substantial additional consumption costs that, ironically, must be imposed on the already-capitated higher-priority individuals if regulators impose price restrictions under pass-through expectations that do not reflect the true extent of additional demand generated.

The 'regulation' examples provide a very cogent illustration why full understanding of both the insurance and health care aspects of health systems must necessarily underpin both their design and their regulation. Unless regulators and policy-makers are fully cognisant of all of the factors that may affect prices in every practice, and recognise that each practice may face a unique set of risks due to its diverse allocation of patient demands and system expectations, it becomes extremely difficult to apply any meaningful form of price regulation. Rather, considerable regulatory risk and potential sector distortion is introduced.

International evidence suggests that regulation of capitated systems is fraught with difficulty. Even in the case of very mature health insurance markets in the United States, the complexity imposed by risk-sharing contracts imposes insurance responsibilities upon individual practices that makes effective price regulation problematic (Hagen, 1999). It is unlikely that the task would be any easier in New Zealand. Furthermore, it is highly unlikely that twenty one individual New Zealand DHBs, more experienced in the provision of hospital care and contract management than health insurance management and regulation, would be as informed and as experienced, and therefore as capable of undertaking the regulatory task as specialist health insurance regulators such as Hagen. Indeed, Ashton (1999) indicates that a fundamental reason why the government chose not to proceed with the creation of competing health care insurers as proposed in the Green and White Paper (Upton, 1990) but instead formed four regional funding monopolies in 1994 related to the potential efficiency and equity consequences and the difficulties of operating and regulating competing private sector health insurers.

4. Conclusion

This paper has illustrated both in theory and in stylised examples that extreme care must be taken when designing capitation contracts in primary care markets that the effects of ‘random’ and ‘predictable’ risk are traded off against each other in a way that ensures that the effects of random risk do not crowd out the desired behavioural changes. Moreover, the paper shows that the design of health systems using capitation contracts must take great care to ensure that other elements in the design of the system do not impede the ability of the capitation contract to achieve its principal aims.

The NZPHCS arrangements provide an example where the effects of random risk appear to have been assumed away, with the consequence that the resulting system, based upon capitation payments to service providers, will achieve neither the desired behavioural changes nor the equity objectives of the strategy. Whilst increased government funding has reduced out-of-pocket costs for patients, total system costs have undoubtedly risen relative to a fee-for-service counterfactual, with the burden of extra costs falling disproportionately on the most frequently ill. Moreover, the vulnerable individuals for whom the system was supposed to provide greater benefits, are also disproportionately affected by other any policy changes that affect the demand for general practitioner services. The costs of ‘random’ risks such as the waiting list cull are certainly being paid at the GP’s surgery, with the sickest paying most of these costs, and the habitually well reaping the benefits.



References

- Arrow, Kenneth. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* 53(5): 941-73.
- Ashton, Toni. 2005. Recent developments in the funding and organisation of the New Zealand Health System. *Australia and New Zealand Health Policy* 2005: 2-9. Available on <http://www.anzhealthpolicy.com/content/2/1/9>.
- Ashton, Toni. 1999. The health reforms: to market and back. In Boston, J., P. Dalziel and S. St John (eds.) *Redesigning the Welfare State in New Zealand: Problems, Policies and Prospects*. Auckland, New Zealand: Oxford University Press.
- Austin, Claire. 2004. *The Public and Private Interface in New Zealand Primary Health Care*. The Royal New Zealand College of General Practitioners. Available on www.rnzcgp.org.nz/PDF/Public_private_0504.pdf
- CBG Health Research Limited. 2004. *National GP Fee Survey*. Wellington, New Zealand: Ministry of Health. Available on [http://www.moh.govt.nz/moh.nsf/0/1BF9A9B04AA23B8ACC256EF4000D3A8A/\\$File/FeeSurveyReportCBG.pdf](http://www.moh.govt.nz/moh.nsf/0/1BF9A9B04AA23B8ACC256EF4000D3A8A/$File/FeeSurveyReportCBG.pdf)
- Coster, Gregor; and Barry Gribben. 1999. *Primary Care Models for Delivering Population-based Health Outcome*. Wellington, New Zealand: National Health Committee.
- Crampton, Peter; Frances Sutton and Jon Foley. 2001. Capitation funding of primary care services: principles and prospects. *New Zealand Medical Journal* 115 (1155).
- Cumming, Jackie. 1999. *Funding Population-Based Primary Care in New Zealand*. Wellington, New Zealand: National Health Committee.
- Danzon, Patricia M. 1997. Tort Liability: a minefield for managed care. *Journal of Legal Studies* 26: 491-519.
- District Health Boards New Zealand. 2006. *Proposed changes to fees framework*. Available on http://www.dhbnz.org.nz/SITE_Default/SIG/FUNDING_ROLLOUT/Default.asp
- Ellis, Randall; and Thomas McGuire. 1986. Provider behaviour under prospective reimbursement: cost-sharing and supply. *Journal of Health Economics* 5(2): 129-51.
- Hagen, Gary. 1999. Shaky Ground: questioning health plan and provider solvency. *Risk Management* 46(11): 42-45.
- Hefford, Martin; Peter Crampton and Jon Foley. 2005. Reducing disparities through primary care reform: the New Zealand experiment. *Health Policy* 72:9-23;
- Howell, Bronwyn. 2005. *Restructuring primary health care markets in New Zealand: from welfare benefits to insurance markets*. *Australia and New Zealand Health Policy* 2:20. Available on <http://www.anzhealthpolicy.com/content/2/1/20>
- Howell, Bronwyn. 2007. *Financial Risk in Primary Health Care Contracting: implications for Sector Structure, Ownership and Outcomes*. Wellington, New Zealand: Institute for the Study of Competition and Regulation.
- Howell, Bronwyn. 2007a. *Contractual Pitfalls in Capitated Primary Health Care: Sharing Random Demand Risk in New Zealand's Strategy*. Wellington, New Zealand: Institute for the Study of Competition and Regulation.
- King, Annette. 2001. *The Primary Health Care Strategy*. Ministry of Health. Available online [http://www.moh.govt.nz/moh.nsf/0/7BAFAD2531E04D92CC2569E600013D04/\\$File/PHCStrat.pdf](http://www.moh.govt.nz/moh.nsf/0/7BAFAD2531E04D92CC2569E600013D04/$File/PHCStrat.pdf)
- King, Annette. 2004. *Memorandum to Cabinet Primary Health Care Strategy- Achieving Low Cost Access*. Wellington, New Zealand: Ministry of Health. Available on http://www.moh.govt.nz/moh.nsf/wpg_index/-Primary+Health+Care+Publications
- Ma, Ching-To Albert; and Michael H. Riordan. 2002. Health Insurance, Moral Hazard and Managed Care. *Journal of Economics and Management Strategy* 11(1): 81-107.
- Malcolm, Lawrence. 1997. GP budget-holding in New Zealand: lessons for Britain and elsewhere? *British Medical Journal* 314(7098): 1890-2.

- McAvoy, Brian; and Gregor Coster. 2005. General practice and the New Zealand health reforms – lessons for Australia? *Australia and New Zealand Health Care Policy* 2:26. Available from: <http://www.anzhealthpolicy.com/content/2/1/26>.
- Milgrom, Paul; and John Roberts. 1992. *Economics, Organization and Management*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Newhouse, Joseph P. 1996. Reimbursing health plans and health providers: efficiency in production versus selection. *Journal of Economic Literature* 34(3):1236-1263.
- Pauly, Mark V. 1968. The economics of moral hazard: comment. *American Economic Review* 58: 531-37.
- Rivers, Patrick Asubonteng; and Kai-Li Tsai. 2003. Managing costs and managing care. *International Journal of Health Care Quality Assurance* 14(6/7):302-307.
- Robinson, James C. 2004. Reinvention of health insurance in the consumer era. *Journal of the American Medical Association* 291(15): 1880-1886.
- Robinson, James C. 2001. Theory and practice in the design of physician payment incentives. *The Milbank Quarterly* 79(2): 149-177.
- Shen, Yu-Chen; and Joshua McFeeters. 2005. *Out-of-pocket health spending between low- and higher-income populations: who is at risk of having high financial burdens?* National Bureau of Economic Research Working Paper 11179. Available on <http://www.nber.org/papers/w11179> .
- Stevenson, Rebecca. 2005. Overworked GPs close books on new PHO patients. *Kapi Mana News* 29/05/05 p1.
- Upton, Simon. 1990. *Your Health and the Public Health: a Statement of Government Policy*. Wellington, New Zealand: Minister of Health.
- Zeckhauser, R. 1970. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory* 2(1): 10-26.