



Reconsidering theories and evidence of supplier induced demand

Professor Jeff Richardson¹
Foundation Director

Stuart Peacock^{2,3}
Senior Scientist (Health Economics)

July, 2006

¹ Centre for Health Economics, Monash University, Australia

² British Columbia Cancer Agency, Vancouver, British Columbia, Canada

³ Department of Health Care and Epidemiology, University of British Columbia, Vancouver, British Columbia, Canada

ACKNOWLEDGEMENTS

The authors would like to acknowledge the assistance of the Commonwealth Department of Health and Ageing and the Victorian Department of Human Services for the provision of the data used in the paper. The research was supported by a NHMRC Project Grant. The views expressed in this paper are those of the authors, and not the funding agency.

Contact Details

For information about the report please contact:

Prof Jeff Richardson, Foundation Director,

Centre for Health Economics, Faculty of Business & Economics, Monash University.

jeff.richardson@buseco.monash.edu.au. Phone: 03 9905 0754

TABLE OF CONTENTS

1. Introduction.....	2
2. SID as a Default Theory.....	3
3. The theory of SID	7
4. Two additional studies.....	11
5. Concluding comments	13
References.....	17
Appendix	17

This paper seeks to make three contributions to the literature on supplier induced demand (SID). They concern, firstly, the criterion for accepting or rejecting SID. Secondly, its theoretical basis and thirdly the results of two Australian studies designed to test the existence of SID in the Australian medical market.

In the first part of this paper (Section 2) it is argued that SID may be described as a 'default theory' which gains its plausibility from its ability to explain and predict service use. From this perspective SID remains an acceptable, instrumental, theory until it is shown to be unnecessary and that the alternative, more orthodox analytical framework has superior explanatory power.

Despite this justification the persuasiveness of the theory is increased if there is a plausible explanation of patient and doctor behaviour which is consistent with it. It is argued that present agency theories are not fully satisfactory. In the second part of the paper two suggestions are made with respect to behaviour based upon the importance of uncertainty facing both a doctor and a patient. It is shown that medical uncertainty (as distinct from asymmetrical information) can explain both patient and doctor behaviour. In contrast with many theorists, SID is shown to be consistent with the ethical behaviour of a professionally and financially motivated doctor optimising demand shift in a competitive and dynamic market.

The first empirical study summarised here used Australian cross sectional data. It is used as a vehicle for commenting on the problem of identification raised by Ramsey and Wassow (1986). The second study draws upon the unique patient and doctor incentives in Australian hospitals which allow a discriminatory test of the relative importance of patients' and doctors' incentives.

The paper draws heavily upon a recently published article by the authors (Richardson and Peacock 2006). It includes arguments which were developed after this publication. These chiefly relate to the default justification of SID in the first section.

Reconsidering theories and evidence of supplier induced demand

1. Introduction

The suggestion that doctors are imperfect agents and can induce demand for health care directly conflicts with the full information and consumer sovereignty assumptions of the orthodox model of demand and supply (Dranove 1988). As a result, supplier induced demand (SID) has been one of the most controversial topics in health economics. Key arguments have been reviewed elsewhere (Folland et al. 1993, Rice and Labelle 1989, Donaldson and Gerrard 1989). Many of the quantitative studies of SID have used cross-sectional data, and examined the effect of the doctor supply on the demand for health care by including doctor supply as an independent variable when estimating the demand equation (Fuchs 1978, Fuchs and Kramer 1972, Richardson 1981, Cromwell and Mitchell 1986, Phelps 1986).

As discussed below, these studies have been criticised on methodological grounds. In order to explain why doctors do not shift demand as far as possible, Evans (Evans 1974) has suggested that doctors have some 'target income' or 'target number of working hours' in mind which they seek to achieve. In countries without controls over medical fees this still implies a conflict with profit maximisation. In principle doctors could shift demand to the limit and reduce the workload to a desired level by increasing fees. To reconcile SID and utility maximising behaviour it has been suggested that SID implies unethical behaviour and that doctors therefore derive disutility from inducing demand (Auster and Oaxaca 1981). For SID to be quantitatively important this would imply widespread and significant malpractice. However the inference that most doctors are consciously providing unnecessary and unethical care lacks *prima facie* validity and conflicts with casual observation. The purpose of this paper is to reconsider some of the theoretical arguments with respect to SID, and to present and review some additional results from the analysis of Australian data.

First, it is argued in Section 2 below that the theory of SID may be described as 'default theory' – one which is 'problem driven' or 'empirically based'. It was proposed to explain observations which were not easy to reconcile with orthodox theory and demand side parameters. This problem is illustrated using Australian data.

In response to the need to provide a theoretical foundation for SID, suggestions have been made with respect to both patient and doctor behaviour. In the second part of the paper (Section 3) these theories are criticised and two suggestions are made relating to patient and doctor behaviour. Both of these are driven by the uncertainty which permeates the health sector with respect to medical outcomes and market behaviour. Empirical studies purporting to measure the extent of SID from cross sectional studies have been criticised because of an alleged failure to identify structural equations. In the third part of the paper this argument is considered and the results of two Australian studies summarised. One of these is based upon cross sectional analysis of the form criticised in the literature. The second draws upon the unique structure of the Australian health system and reports the results of the 'natural experiment' which arises from the funding of public and private hospital patients.

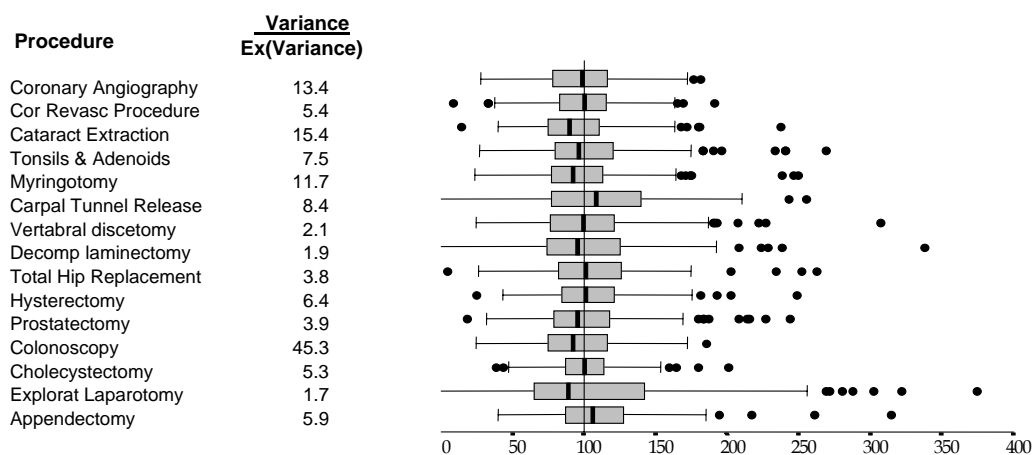
2. SID as a Default Theory

A theoretical model of the demand and supply of medical services should be capable of providing an explanation of observed patterns of service use, both cross-sectionally and through time. Six sets of Australian data are discussed below, which indicate the utilisation patterns that must be explained by our theory. For four of the six sets of data only the correlation between doctor supply and utilisation is available. For well known reasons correlation does not indicate causation. However it does indicate the existence of a relationship which requires explanation. This may be, firstly, that there is a causal sequence from demand to supply (when demand increases supply responds). Secondly, there may be an unobserved variable such as queuing which explains why demand adjusts to supply. Finally there may be a causal sequences from supply to demand that is, supplier induced demand. While correlation does not identify the direction of causation the elimination of the first possibilities is strong evidence of the existence of the third causal mechanism, that is, supplier induced demand may be viewed as default theory to explain empirical observations when other explanations of utilisation appear to be improbable. Historically it was this argument which was used in the seminal article on SID by Evans (1974) in the context of service use across British Columbia in Canada.

The first set of observations relates to fifteen well defined hospital procedures and the variation in their use across the state of Victoria. The subsequent datasets consist of GP and specialist supply and service use. A final choice between SID and orthodox theory should be based to a very large extent upon the explanatory power of the rival theories for all of these patterns.

The first data set was used to plot differences in the two year utilisation rates per 1,000 population in the statistical sub-divisions in the State of Victoria. The results shown in figure 1 are standardised for age and sex and normalised so that the expected use per 1,000 population is set equal to 100. Points in the box plots show actual utilisation relative to this norm. The figure identifies a 4-6 fold variation in the use of different procedures. This cannot be attributed to random variation. When the variance in the utilisation rates for each procedure is divided by the variance predicted from the age/sex composition of each area,ⁱ the ratio shown to the left of each plot varies from 1.7 for exploratory laparotomy to 45.3 for colonoscopy.

Figure 1: Standardised Rate Ratios for Various Operations in the Statistical Local Areas in Victoria, compared with the Rate Ratios for all Victoria



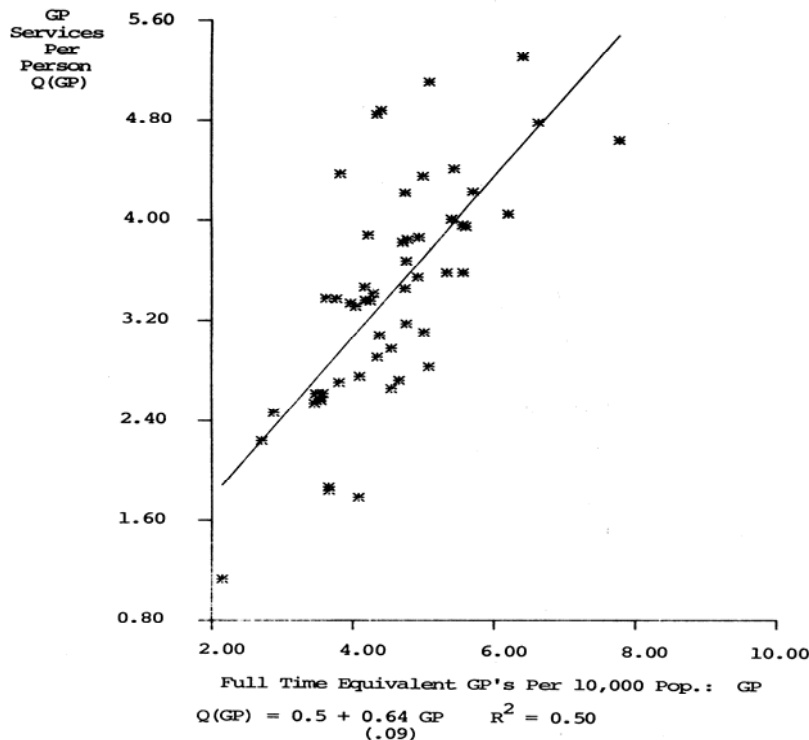
Source: Richardson (Richardson 1998)

Note: Bold line indicates the Median value; dark rectangles represent 25th and 75th percentiles for statistical local areas, standardised to Victorian state ratio = 100. Extreme values greater than three times 50th – 75th and 25th – 50th percentile intervals are recorded as separate points
See Appendix for explanation of small area units of analysis

Small area variation such as these have been widely reported (McGuire 2000). Four conclusions are self evident from these data. First, the variation to be explained is quantitatively very large and well outside the magnitudes associated with price and income effects. Secondly, it is implausible to posit that they are driven by medical need. The populations in the different small areas are relatively homogeneous and the reported results have been age-sex standardised. Thirdly, it is equally implausible that the results are driven by consumer tastes. Most of the procedures are unpleasant and patients would need persuasion of their need. Fourthly, the majority of the procedures are complex and their outcomes sufficiently uncertain that the decision to undergo the procedure would necessarily require an expert assessment of the likely benefits. Conversely, the expert-agents, not the patient's assessment will be used in a comparison of costs and benefits.

Figures 2 and 3 constructed from the next data sets, indicate that there has been, and remains, a very close relationship between the geographic availability of GPs across Australia and the use of their services. As there is significant border crossing, especially between statistical sub-divisions (the unit of analysis in figure 3) the relationship is not a necessary one (see appendix for small area definitions). In principle, local supply could be unrelated to local demand, with border crossing and variable work loads accounting for the discrepancy. The two figures do not demonstrate a causal relationship between GP supply and the demand for their services for at least two reasons. First, increased border crossing would be associated with increased time costs. Consequently, poorly supplied regions would have lower demand because of these costs. However, from the perspective of a government interested in the creation of equal access or in the allocation of a health budget, the distinction between SID, as usually conceived, and a supplier induced variation in time costs which explains demand, is of very little interest. The close correlation suggests, as a minimum, that SID could be accepted by government economists as an instrumental theory; that is, as a theory which is useful for prediction and policy if not for the description of individual behaviour. Importantly, however, border crossing was not a significant determinant of demand in Richardson's 1981 study (Richardson 1981), and in the recent re-evaluation of Australian evidence (Peacock and Richardson 2005, Richardson and Peacock 1999).

Figure 2: GP Supply vs. Use by Statistical Division, Australia, 1976



Source: Richardson & Deeble (Richardson and Deeble 1982)

Note: See Appendix for explanation of small area units of analysis

Figure 3: GP supply and use, by Statistical Sub-Division 1996/97

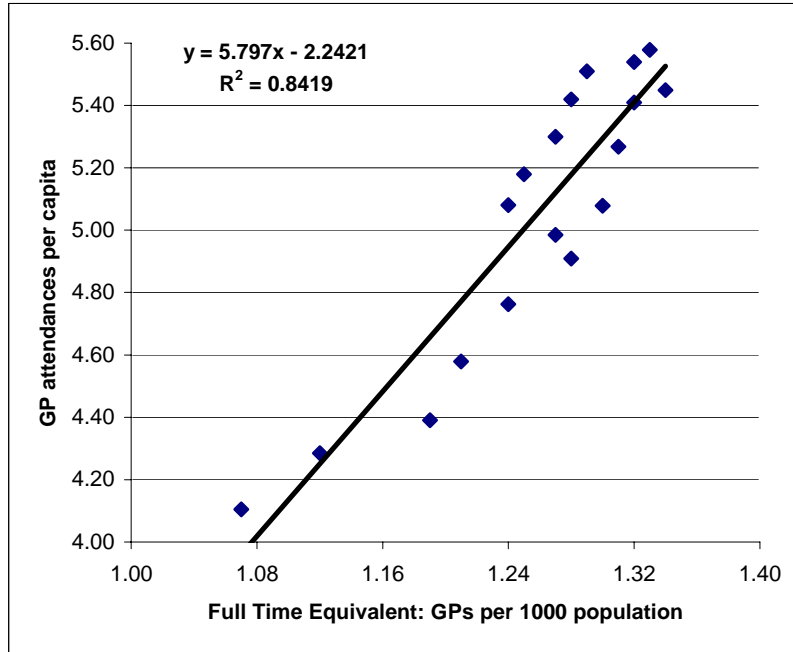


Note: See Appendix for explanation of small area units of analysis

The significance of these two figures is twofold. First, they demonstrate the possibility of SID: such a correlation is a necessary, but not sufficient condition for the existence of SID. Second, it is possible that GPs locate their practices in areas of high autonomous demand. It is for this reason that a serious statistical analysis of cross-sectional data must attempt to take account of reverse causation and endogenise the GP supply (see section below). The importance of reverse causation could, however, be overstated. Age/sex standardisation does not reduce the variation in either the 1976 or 1996 data significantly. Price, income and socio-economic variation are also insufficient to explain a significant part of the variation. Data in figure 2 were modelled in Richardson (Richardson 1981) and the data in figure 3 by Peacock and Richardson (Peacock and Richardson 2005). The later results are reported below.

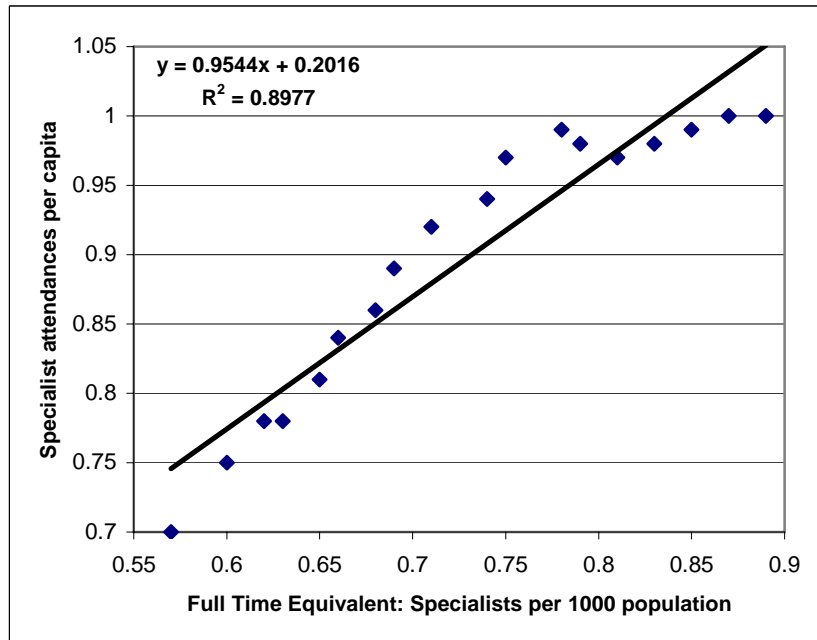
The most persuasive correlational data, are those presented in figures 4 and 5. They indicate that over time there has been a nearly perfect correlation between the growth of general practitioners and specialists, and the use of their services, both in term of consultations (figures 4 and 5) and diagnostic services (figure 6). Unless it is postulated that there has been a permanent excess demand for these services, then causation must run from supply to demand. The apparent impact of supply upon demand is most clearly illustrated following the cessation of the growth in the GP supply in 1995 and the corresponding cessation in the growth of service use. Observations below the trend line represent years in which the GP supply was expanding. The seven observations above the line represent years when the supply was contracting. The difference does suggest that the latter years may have experienced excess demand, possibly as a result of an upward ratcheting of expectations. However over the full time period the magnitude of the change cannot be attributed to increases in per capita income or to the small changes in patient net payments which occurred. The remaining demand side variable, time cost, may have contributed to the correlation during the contractionary years but is unlikely to have been the chief causal agent overall, and particularly during the years of growing GP supply, 1984-1995. The near consensus during this period was that the medical market was characterised by excess supply, not excel demand (Scotton 1984). It was for this reason that governments sought, successfully, to reduce the number of GPs relative to population. **

Figure 4 GP supply and use 1985-2002



Source: Commonwealth Department of Health and Family Services, Data file (Commonwealth Department of Health and Family Services 1998).

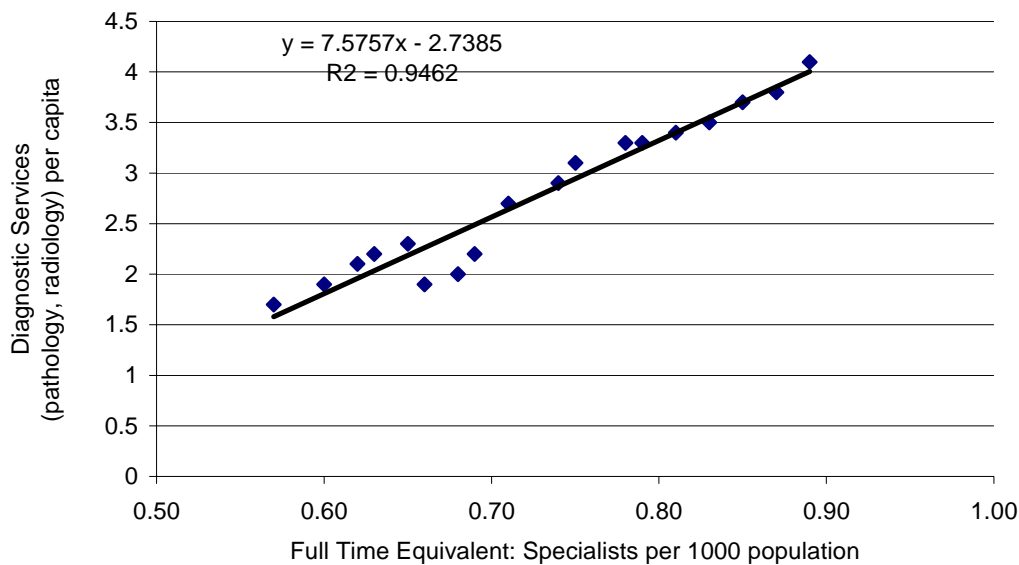
Figure 5: Specialist supply and use 1985-2002



Source: Commonwealth Department of Health and Family Services, Data file (Commonwealth Department of Health and Family Services 1998).

Note: Specialist supply excludes Pathology and Radiology

Figure 6: Specialist supply and use of diagnostic services 1995-2002



Source: Commonwealth Department of Health and Family Services, Data file (Commonwealth Department of Health and Family Services 1998).

Note: Specialist supply excludes Pathology and Radiology

As noted, correlational evidence is weak when the direction of the causation is in doubt, and when one or more other variables contaminate the correlation because they are also correlated with the variables of interest. However, correlational evidence is far more powerful when, as in the present case, reverse causation is improbable, and when other variables cannot be identified which confound the interpretation of the correlation. In particular, the doctor supply in Australia is largely determined by exogenous national and state policies relating to medical school places. Further, there is no other variable which is known to have changed to such an extent that it is able to explain such a large change in the use of GP services, which have remained relatively homogeneous despite the technological advances in specialist services.

The single caveat is that data are not available with respect to time costs, either in terms of days until a doctor may be seen or hours waiting in a doctor's office. With respect to GPs, the casual evidence referred to above does not suggest significant time costs until quite recently when the supply of GPs fell relative to population. This caveat is potentially important if the social or government objective in the health sector is the maximisation of utility as defined in welfarism.

3. The theory of SID

While the theory of SID may have evolved from a pragmatic explanation of observed variation in services use, it has also been necessary to provide an explanation of patient and doctor behaviours consistent with the theory. More specifically, it has been necessary to explain why patients allow their revealed preferences to be determined, at least in part, by doctors and why doctors do not shift demand to its technical limit, control excess (induced) demand by raising fees and thereby maximise profit and personal utility.

With respect to the first question it is generally acknowledged that patients face asymmetrical information in the agency relationship with the doctor – patients rarely possess the technical

knowledge necessary to determine the expected benefits from their health care consumption choices. (There are undoubtedly exceptions to this generalisation in the case of very simple services or in the regular regime of treatments needed by patients with a chronic disease.) We suggest, however, that this explanation is not complete or necessarily the most important element in the abdication of decision making by patients. As discussed below, medical decision making is characterised by uncertainty and, particularly at the level of the individual, decision making requires not just information, but judgement. Even with well informed patients there is an asymmetrical ability to assess and synthesise information and to draw upon past experience to help reach a conclusion with respect to the optimal course of action. Doctors have clinical experiences which patients do not. Doctors have accepted the role as decision maker whereas patients, recognising their inexperience, rely upon the doctors' judgement. In sum, even in the absence of an asymmetry in the technical information available, a well educated and normally 'empowered' individual is likely to ask for the doctor's judgement and accept that the doctor's experience makes this judgement better than the patient's judgement. At the level of the patient SID is little more than the theory that patients generally trust their doctor's judgement more than their own.

Discussion of doctor behaviour in the context of SID has been clouded by the debate over what is the medically, and ethically, correct level of service provision in health care. SID has often been characterised a form of unethical behaviour, and that it is recognition of their own unethical behaviour that has restrained doctors from fully exploiting SID. However unless SID was quantitatively small this assumption implies widespread unethical behaviour by doctors, and with respect to a significant number of patients and treatments. The inference of such misconduct in a profession which is widely regarded as being highly professional appears improbable and there is no supporting evidence.

However SID need not imply that doctors are behaving in an unethical manner. As noted by Dranove under certain conditions a doctor will have an incentive "to recommend treatments whose costs outweigh their medical benefits" (Dranove 1988, p281). This outcome may be the result of supply side moral hazard. Doctors in their role as agents may deliver services to patients until the marginal benefit was zero despite a positive cost to society. This behaviour may be characterised as being socially inefficient, rather than unethical.

These ethics based theories imply an understanding of the benefits of individual marginal services which is not supported by the literature. In the last two decades it has become increasingly apparent that there has been no well defined level of services which has been accepted as medically and ethically correct (Maynard and Bloor 1998, Mooney et al. 1986, Hurley et al. 1997). The evidence based medicine movement is a reaction to this. Wennberg summarised this characteristic of medical care when he notes that the validity of the many specific theories physicians hold on appropriate practices, are now recognised as problematic and professional uncertainty rather than consensus about the scientific basis of clinical practice is emerging as the dominating reality (Wennberg 1988, p101). In part, this uncertainty is a function of the small number of services that have been evaluated even for clinical efficacy. One OECD study suggests that only 20 percent of procedures in common use have been evaluated (Oxley and MacFarlan 1994). Rather than reflecting an established set of responses to well-defined indicators, practice patterns appear to vary with the myriad variables that influence clinical decision-making: training, peer behaviour, conference attendance, personal temperament, personal experience, financial rewards, and, most importantly here, time and infrastructural capacity to undertake more or less intensive investigative and therapeutic work.

Based on the variation in the patterns of care such as shown in figure 1, and numerous similar studies, Wennberg's main conclusion is persuasive: "The evidence from small area analysis, from the critical appraisal of strengths and weaknesses of the scientific basis of medicine and

the failure of expert panels to reach consensus on appropriate practice build a consistent and strong case against the rational agency hypothesis and the associated assumptions about the nature of demand in medical markets” (Wennberg 1988, p 100). This suggests a more complex theory of the demand and supply of medical services than proposed in the orthodox model. First, patients face asymmetric information and an asymmetrical capacity to evaluate in the face of uncertainty. But, secondly, and following Wennberg, uncertainty may characterise not just patient decision-making but also the decisions made by doctors. Under these circumstances doctors may understandably believe that ‘more care is better than less’, a belief that would justify the creation of demand to the point where they reached their leisure constraint, while simultaneously believing that the induced demand represented better quality and ethically commendable care.

If demand inducement is not unethical, then there is the unexplained question of why doctors do not shift the demand curve to its limit and control excess demand by increasing their fees. Not doing this conflicts with the profit maximisation assumption of the orthodox model of the supply of health care by doctors. However, to suggest that doctors’ sole objective is the pursuit of income or profit seems implausible. Evans first suggested that doctors do not induce demand to its technical limit because they have some ‘target income’ or ‘target number of working hours’ in mind which they seek to achieve (Evans 1974). That is, doctors, like all other economic actors, face an income-leisure trade-off decision.

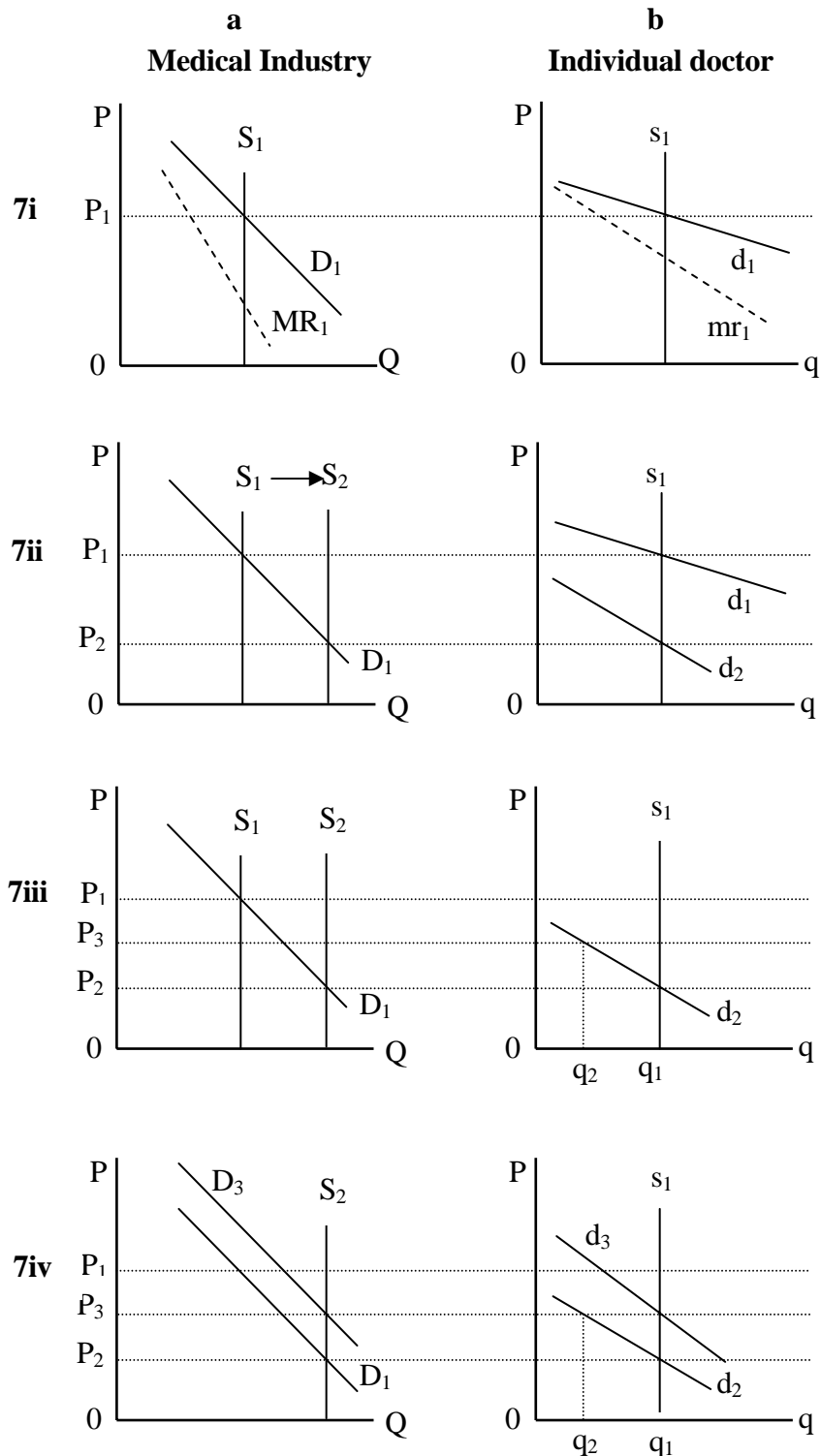
If the profit maximisation assumption is relaxed, so that doctors are allowed to have a target number of working hours, and three further weak (and plausible) assumptions are added to the orthodox model of utility maximisation, then doctor’s behaviour and motivation associated with SID is easily explained. These four assumptions are that:

- (i) medical prices are ‘sticky’ in the sense that doctors are reluctant to significantly change them in the short-run;
- (ii) *individual* doctors face relatively elastic demand curves;
- (iii) doctors have limited knowledge of overall industry conditions; and,
- (iv) there is a target number of working hours which the doctor, for personal and professional reasons, seeks to achieve.

With these assumptions, incomplete SID is a likely consequence of the dynamics of market adjustment following an increase in the supply of doctors. This is illustrated in figure 7. In the initial equilibrium – figure 7(i)a – there is a relatively inelastic demand for the industry curve and a completely inelastic supply set by the number of doctors and their target working week. With either a competitive or monopolistic model the equilibrium, profit maximising price P_1 will clear the market. The individual doctor faces a price elastic demand curve as patients can switch relatively easily between doctors. The curve is not completely elastic because of patient loyalty to ‘their’ doctor. The individual doctor is also in equilibrium at P_1 (figure 7(i)b).

Figure 7(ii) shows the final equilibrium in a market with no demand shift following an increase in supply from S_1 to S_2 . Industry supply and demand result in market clearing at price P_2 . For the individual doctor, the increase in supply results in a reduction in demand from d_1 to d_2 (figure 7(ii)b). However this second equilibrium is never reached. Figure 7(iii) illustrates the dynamics of the adjustment process. In the short-run ‘sticky prices’ fall only to P_3 (figure 7(iii)a) and at this price there is excess supply. The individual doctor will experience a significant reduction in demand from q_1 to q_2 (figure 7(iii)b). In response to this the doctor can shift demand from d_2 to d_3 (figure 7(iv)) at which point both industry and doctor are in a new equilibrium.

Figure 7: Adjustment to a Change in the Supply of Medical Services



Notes: ¹ Assuming a Poisson distribution for the use of health services in each age cohort.

Limited demand shift has occurred and the doctor is at a new profit maximising equilibrium consistent with the target working week. Additional demand shift at this price would encounter the leisure constraint. If the doctor were to experimentally increase price, demand would fall significantly. Unlike the previous decrease in demand the doctor could attribute this directly to his or her own pricing decision which would appear to be unprofitable. Because of poor market information and non collusive behaviour, this would induce the doctor to accept P_3 as a final equilibrium price.

The behaviour postulated here would be less likely to occur if the doctor's only motivation was the maximisation of profit. Repeated experimentation with incremental increases in price and further demand shift might increase profit. However this behaviour might well be regarded as unethical. The experimental increases in price would intentionally reduce the demand of less well off patients and appear unprofessional in a way that increasing the amount of attention given to a patient does not. In sum, the behaviour postulated here is that the majority of doctors have some interest in short-run profit maximisation, but are also motivated by the achievement of professional objectives through the provision of what they judge to be useful services.

The behaviour described in figure 7 is consistent with weak profit maximisation and demand shift. However it is only one possible outcome from the exploration of the dynamics of the market. Demand shift with rising prices could similarly be explained if falling demand per doctor resulted in an increased price to maintain income, and demand shift to achieve the target working week. The general point is that there are a number of explanations for doctor behaviour that are consistent with demand shift if the dynamics of the market are explored in combination with the assumption of weak profit maximisation, belief in the efficacy of the services provided, and a professional commitment to ethical behaviour.

4. Two additional studies

The identification problem: In a recent study we re-examined criticisms of cross-sectional methods used to test for supplier induced demand (SID), and re-evaluated the empirical evidence using data from Australian medical services (Peacock and Richardson 2005, Richardson and Peacock 1999). We briefly summarise our findings here, because they have important implications for the SID debate.

Cross-sectional studies that have included the doctor supply as an independent variable in the demand equation to test for the presence of SID, have been criticised on two grounds. The first, and most important, criticism is that under the orthodox specification of demand and supply equations the SID effect cannot be econometrically identified if the *supply of health care services* is added into the demand equation (Auster and Oaxaca 1981). This has commonly been interpreted in the literature to mean that the inclusion of the *doctor supply* in the demand equation leads to an identification problem (Richardson and Peacock 1999). The second criticism is that early empirical studies of SID fail more recently developed diagnostic tests, and produce artefactual findings due to model misspecification (Ramsey and Wassow 1986).

Auster and Oaxaca's original criticism remains valid (Auster and Oaxaca 1981). However, in our (and others') empirical work, including Fuchs' original econometric modelling (Fuchs 1978, Fuchs and Kramer 1972) this does not lead to an identification problem, because the doctor supply enters the demand equation as an independent variable, not as the supply of services (Peacock and Richardson 2005, Richardson and Peacock 1999). The doctor supply variable is stochastic, and it depends upon a variety of variables including the desirability of the location. However, Auster and Oaxaca's original criticism has commonly been interpreted in the literature to mean that the inclusion of the doctor supply in the demand equation leads to an identification problem. This interpretation is invalid. In two-stage least squares regression the reduced form equation for each of the endogenous variables has the same exogenous variables. There is no reason why a change in the structural equations must necessarily change the reduced form equations (Peacock and Richardson 2005, Richardson and Peacock 1999). Ramsey and Wassow's criticism did not apply to our analysis (Ramsey and Wassow 1986). Our findings showed that demand equations which *did not include the doctor supply* were misspecified (Peacock and Richardson 2005, Richardson and Peacock 1999).

Cross sectional Australian results: Empirical evidence from the re-evaluation of Australian medical services data supported the notion of SID. Demand and supply equations were well specified and had very high explanatory power. The demand equation was identified, and the desirability of a location was an important predictor of the doctor supply. Results showed an average price elasticity of demand of 0.22, and an average elasticity of demand with respect to the doctor supply of 0.46, with the impact of SID becoming stronger as the doctor supply rose. The conclusion we can draw from this re-evaluation, is that two of the main criticisms of the empirical evidence supporting the SID hypothesis have been inappropriately levelled at the methods used. More importantly, SID provided a satisfactory, and robust, explanation of the empirical data on the demand for medical services in Australia.

A discriminating test: Australia provides a unique laboratory for the examination of the relative importance of price and SID effects on patient demand. The peculiar financing of Australian health services results in a public sector in which hospital patients are treated without cost, but a private sector in which, following the purchase of private health insurance, the patient is still left with significant out of pocket expenses. In a simple market equilibrium public demand per capita would be expected to exceed private demand per capita. However, incentives facing doctors also differ between public and private sectors. In the public sector there is no financial benefit from the treatment of additional patients. In the private sector a full fee is earned. There is therefore no financial incentive for doctors to increase demand in the public sector, but there is a strong financial incentive to increase demand in the private sector.

In a recent study Richardson and Robertson (Richardson et al. 2002, Robertson and Richardson 2000) examined the treatment of patients after an emergency admission with a heart attack (acute myocardial infarction [AMI]). Various treatments are possible for AMI. The most expensive and recent of these include the diagnostic test, angiography, and the procedures collectively known as 'revascularisation', that is, coronary artery bypass surgery, balloon angioplasty and stenting. Each of these four procedures attracts a significant fee in the private sector.

For the two years of the study 100 percent of the procedures delivered in the state were analysed. The likelihood of a procedure for private patients in both private and public hospitals was divided by the likelihood of these procedures for private patients in a public hospital. The results, reported in Table 1, indicate a significantly greater likelihood of a procedure in the fourteen days following admission for all categories of private patients. The discrepancy varied from 57 percent in the case of female private patients in public hospitals receiving angiography to 286 percent for female private patients in private hospitals receiving a revascularisation procedure. In the subsequent twelve months these discrepancies decreased slightly but for patients initially admitted to a private hospital the discrepancy remained over 100 percent for all patients. The inescapable conclusion appears to be that these patterns were driven by physician judgement rather than patient preference; that is, that doctors, not patients, determined the use of services.

Table 1: Ratio of likelihood of a procedure for private to public patients in private and public hospitals, 1995/97

	Private hospital patients : public hospital patients to		Private patients in public hospitals : public patients to	
	Angiography	Revascularisation	Angiography	Revascularisation
Within 14 days				
Men	2.20	3.43	1.77	1.53
Women	2.27	3.86	1.57	1.81
Within 3 months				
Men	2.24	3.43	1.53	1.23
Women	2.28	3.34	1.49	1.32
Within 12 months				
Men	2.16	2.89	1.42	0.97
Women	2.22	2.84	1.48	1.10

Source: Richardson et al. 2002, Robertson and Richardson 2000) Original Data Source: Victorian Inpatient Minimum Dataset

5. Concluding comments

There have been three parts to this paper. In the first we have reported data to demonstrate that in the Australian health sector, as elsewhere, one of the most important facts requiring explanation is the enormous variation in the rate of service use. Perhaps the most persuasive argument for SID is that it explains much of this variation. In contrast the usual determinants of demand do not vary significantly and cannot provide such an explanation.

In this context the SID literature has not been even handed in its treatment of evidence and theory. Proponents of SID have been asked for a satisfactory theory of consumer and provider behaviour. Without this it is plausible to argue that, by default, orthodox theory should be retained. However it is equally plausible to argue that if this orthodox theory does not explain the large facts of the medical market then, by default, SID should be embraced. While most would agree that, in principle, both criteria should be of importance this is not reflected in the literature where the onus of proof is primarily required from proponents of SID.

In the second part of the paper we have argued that the most plausible explanation of patient and provider behaviour arises from the uncertainty which permeates the health sector. At the theoretical level 'SID' may be no more complex than the hypothesis that patients follow their doctor's advice and that doctors believe their services to be beneficial and that it is professionally acceptable to provide more care to patients when they have the capacity to do so.

Patients transfer decision making to their doctor-agents because of uncertainty. This is not only due to the asymmetrical information available to patients and doctors but also to the asymmetrical capacity to draw upon past experience to *assess a potentially unique set of symptoms; that is*, patients are disadvantaged by their lack of experience.

The explanation for doctors' not fully exploiting SID can be found in the dynamics of market adjustment when doctors have limited market information, when there is no collusion and when there is satisfaction from the achievement of professional goals.

In the final part of the paper we have provided further evidence of SID using two Australian studies. For the first, cross section data were used. We argue that two of the main criticisms of the empirical evidence supporting the SID hypothesis have also been inappropriately levelled

at the cross-sectional methods used to test and quantify SID. When these methods are applied in Australia SID provides a satisfactory, and robust, explanation of the observed variation in services per capita. Importantly, omission of the stochastic variable for the doctor supply results in the misspecification of the demand equation.

Australia provides a unique laboratory for the examination of the relative importance of price and SID effects on patient demand. Small area data shows that private sector patients are between 2 and 3 times more likely to receive certain cardiovascular procedures than public sector patients. Under the orthodox model we would expect to see the opposite, i.e. much greater use of services or significant queues, in the public system where care is free at the point of consumption. This is not the case. It seems an inescapable conclusion that these utilisation patterns are driven primarily by physicians and not by patients.

We conclude, therefore, that once the uncertainty of medical decision-making, the complexity of medical judgements and dynamic adjustment with imperfect market information are taken into account, SID is based upon a more plausible theory of patient and doctor behaviour than the orthodox model. More importantly, SID provides a satisfactory explanation of the observed pattern and change in the demand for Australian medical services. Variations in other observed variables appear to be incapable of providing this explanation within the orthodox framework.

REFERENCES

- Auster R, Oaxaca R. 1981. Identification of supplier induced demand in health care sector. *Journal of Human Resources* 16, 327–42.
- Australian Bureau of Statistics (ABS). 1999. *Australian Standard Geographical Classification (ASGC) 1999*. Canberra: Australian Bureau of Statistics.
- Commonwealth Department of Health and Family Services. 1998. *Medicare statistics: 1984/85 to March quarter 1998*. Canberra: Commonwealth Department of Health and Family Services.
- Cromwell J, Mitchell JB. 1986. Physician-induced demand for surgery. *Journal of Health Economics* 5, 293-313.
- Dranove D. 1988. Demand inducement and the physician/patient relationship. *Economic Enquiry* 26, 281–98.
- Doessel DP. 1998. Is an increased medical workforce a “problem” in the health sector?: theory and evidence. In: Harris, A. (Ed.), *Economics and Health 1997: Proceedings of the Nineteenth Australian Conference of Health Economists*. Australian Studies in Health Services Administration University of New South Wales.
- Donaldson C, Gerrard K. 1989. Paying general practitioners: shedding light on the review of health services. *Journal of the Royal College of General Practitioners* 39, 114-117.
- Evans RG. 1974. Supplier-induced demand: some empirical evidence and implications. In Perlman, M. (Ed.), *The Economics of Health and Medical Care*. New York: John Wiley and Sons.
- Folland S, Goodman AC, Stano M. 1993. *The Economics of Health and Health Care*. MacMillan: New York.
- Fuchs V. 1978. The supply of surgeons and the demand for operations. *Journal of Human Resources* 13 (suppl.), 35-56.
- Fuchs VR, Kramer MJ. 1972. *Determinants of Expenditures for Physicians’ Services in the United States, 1948–68*, National Bureau of Economic Research Occasional Paper no. 117, DHEW Publication (HSM), National Center for Health Services Research and Development.
- Hurley J, Birch S, Stoddart G et al. 1997. Medical necessity, benefit and resource allocation in health care. *Journal of Health Services Research and Policy* 2:223-230.
- Maynard A, Bloor K. 1998. *Our Certain Fate: Rationing in Health Care*. London: Office of Health Economics.
- McGuire T. 2000. Physician Agency, in (eds) A.J. Culyer and J.P. Newhouse, *Handbook of Health Economics* Volume 1B, Chapter 9, Elsevier: North Holland.
- Mooney G, Russell E, Weir R. 1986. *Choices for Health Care: a Practical Introduction to the Economics of Health Provision*. London: Macmillan.
- Oxley H, MacFarlan M. 1994. *Health care reform: controlling spending and increasing efficiency*. Economics Department Working Paper No 149, OECD, Paris.
- Peacock SJ, Richardson JRJ. Supplier induced demand re-examined and re-evaluated. *Health Economics*, (revised and resubmitted). See also (Richardson and Peacock 1999).
- Phelps C. 1986. Induced demand – can we ever know its extent? *Journal of Health Economics* 5, 355-365.

- Ramsey JB, Wassow B. 1986. Supplier induced demand for physician services: theoretical anomaly or statistical artefact? An econometric evaluation of some important models in physician service markets. In: Basmann, R., Rhodes, G. (Eds.). *Advances in Econometrics*, JAI Press, Greenwich, 49–77.
- Rice T, Labelle RJ. 1989. Do physicians induce demand for medical services? *Journal of Health Politics, Policy, and Law* 14, 587-600.
- Richardson J. 1981. The inducement hypothesis: that doctors generate demand for their own services. In: van der Gaag, J., Perlman, M. (Eds.), *Health, Economics and Health Economics*. Amsterdam: North-Holland Publishing Co, 189–214.
- Richardson J. 1998. The health care financing debate. In: Mooney, G., Scotton, R.B. (Eds.), *Economics and Australian Health Policy*. Sydney: Allen & Unwin.
- Richardson J, Deeble J. 1982. *Statistics of private medical services in Australia 1976*. Health Research Project, Technical Paper No 1, Canberra: Australian National University Printing Service.
- Richardson J, Peacock S. 1999. *Supplier induced demand reconsidered*, Working Paper No. 81, Centre for Health Program Evaluation, Monash University, Melbourne.
- Richardson J, Peacock S, (2006). S Reconsidering theories and evidence of supplier induced demand, *Applied Health Economics and Health Policy*, vol 5 no 2.
- Richardson J, Robertson I, Hobbs, M et al. 2002. The impact of new technology on the treatment and outcomes of acute myocardial infarction in Australia, in (Eds.) M.B. McClellan & D.P. Kessler, *Technological Change in Health Care: A Global Analysis*, The University of Michigan Press, Ann Arbor, Chapter 7.
- Robertson I, Richardson J. 2000. The effect of funding upon hospital treatment: the case of coronary angiography and coronary artery revascularisation procedures following acute myocardial infarction. *Medical Journal of Australia*. 173, 291-295.
- Scotton R.B. 1984. Medical Manpower, in (ed) M. Tatchell, *Perspectives on Health Policy*, Health Economics Research Unit, The Australian National University, Canberra, pp 175-185.
- Wennberg J. 1988. Improving the medical decision making process. *Health Affairs* 7, 99-106.

Small Area Definitions

Small area data used in empirical analysis is based on the hierarchical structure of the Australian Standard Geographical Classification (ABS 1999). In non-census years the classification consists of Statistical Local Areas (SLAs), Statistical Sub-Divisions (SSDs), Statistical Divisions (SDs), and States/Territories. Under the hierarchical structure SLAs are aggregated to form SSDs, SSDs are aggregated to form SDs, and SDs aggregate into States and Territories. These spatial units cover all of Australia without gaps or overlaps. As at 1999, there were 1,331 SLAs, 194 SSDs, and 66 SDs covering mainland and offshore Australian States and Territories. SLAs and SSDs are based on defining regions that show social and economic homogeneity by identifiable links between inhabitants, and on local government boundaries. SDs also maintain this basis, but in addition the capital city of each State/Territory is defined as a single SD.
