Enterprise liability, risk pooling and diagnostic care¹

Draft 10 – February 2010

¹ We are grateful to two anonymous referees and the editors for valuable comments. Remaining errors are ours.

Abstract

The safety of patients is an important responsibility of health care providers, and significant compensation costs may arise if providers are negligent. A widely debated option involves liability for such compensation being placed with the hospital rather than the individual clinician, a system known as “enterprise liability”. In the US, partial adoption of enterprise liability, and proposals for its universal introduction, have accompanied high profile “malpractice insurance crises” in the last two decades. Hospitals in England and Wales have been subject to this system since 1990, and risk pooling arrangements have emerged subsequently allowing hospitals to transfer their liability risk to an agency known as the NHS Litigation Authority. We explore some of the mechanisms used by this agency to provide hospital management with financial incentives to take care. We estimate the influence of these arrangements on the use of diagnostic imaging tests within hospitals, using a panel data set covering the period 2000-2004, during which period a policy shift took place leading to a form of “natural experiment”. Our results suggest that the use of diagnostic tests did not respond to the incentives created during this period. We speculate that certain types of patient care activity, including the use of diagnostic tests, may be less responsive to incentives placed at the level of the hospital by comparison with incentives placed at the level of the clinician. Our findings may have implications for jurisdictions contemplating a move to enterprise liability, as well as wider implications for public sector organizations faced with financial incentives to improve service quality.

1. Introduction

Health care providers owe a duty of care to those they serve, and a breach of that duty can be used as a basis for compensation claims. In many countries, including England and Wales, tort liability is the way the law arranges for those who believe they have been harmed by someone to recover a sum of money in compensation for the harm caused (see Fenn et al, 2004). Both the merits of the claim, and the amount of compensation, are determined through a legal process which is underpinned by the ultimate option of trial in a court before a judge. These procedures are well-established for many types of harm including medical negligence. Patients whose claims are successful (in pre-trial settlement or at trial) receive compensation for economic loss, future health care costs, and for pain and suffering.²

Although this broad approach underlies all tort law, a number of significant variations exist in the approaches that health care systems take towards implementing liability rules. One important distinction relates to *who* is the subject of the claim: the clinician (who made the error) or the hospital for whom s/he works?³ In England and Wales, for example, since 1990, the hospital has accepted responsibility for the errors of its clinicians, under a system known as “enterprise liability” – or “NHS indemnity” as it was called in England and Wales. In the US, by way of contrast, it is typically clinicians who are the object of medical negligence suits though, in a number of States, health maintenance organizations (HMOs) and managed care organizations (MCOs) have decided not to contest judgements holding them vicariously liable for their staff’s errors, thereby creating *de facto* enterprise liability – what Leone, Jr (2006) refers to as the “virtual reality of enterprise liability”⁴. An ongoing policy debate in the US has sought to encourage wider adoption of enterprise liability but this has yet to be resolved.⁵

² Other means of redress include making a complaint; complaints procedures normally cover a wider range of circumstances than claims for patient compensation. See Wallis and Mulcahy (2001).

³ Another distinction is between *strict* liability (where the injurer is liable for damages whatever the care supplied) and *negligence* liability (where the injurer is only liable if the level of care supplied is found to be below that required by the law). Our focus is on the latter liability rule: this characterises NHS liability and, indeed, that in operation in many other health care systems, including the US.

⁴ Indeed this was the process by which NHS indemnity emerged in the UK – the NHS autonomously decided that it would not assert its right to recover damages from its employees after meeting its vicarious liability for their errors. The management of NHS hospitals were subsequently bound by this agreement.

⁵ The earliest proposals for a system resembling enterprise liability in the US date back to Steves (1975). As Bovbjerg and Berenson (2006) document, this and subsequent suggestions have typically reacted to crises in the market for clinicians’ medical negligence insurance (though Baker (2005) contests the significance of such crises). For examples of more proposals for the US, see Weiler (1991), Abraham and Weiler (1994)

The theoretical properties of liability rules themselves are well understood. In principle, making the ‘tortfeasor’ (i.e. the party responsible for the accident) liable for the damages s/he has caused induces optimal levels of care, in the sense that the marginal benefit of reduced accident probabilities is equated with the marginal cost of the care (see Shavell, 1987). Yet such optimal deterrence can quite readily be undermined (again, Shavell, 1987, provides a comprehensive treatment). For example, anticipated court error at the trial stage can send inaccurate signals about expected damages which, in turn, affect potential tortfeasors’ marginal calculations. Or situations where the potential tortfeasor has insufficient resources to meet liabilities can also dilute incentives for care since, above a certain level of damage, s/he cannot pay the expected damages (i.e. s/he is ‘judgement-proof’). One solution, relevant to our paper as we explain below, sees the tortfeasor purchase insurance against future liability: while reducing the prospect of judgement-proofness, incentives for care can be undermined if the insurance is so complete as to shield the tortfeasor from all consequences of any harm s/he may cause.

For these reasons (and others), the actual effects of liability rules require empirical study, but the literature has found it difficult to reach unambiguous conclusions about the role of tort liability in encouraging care in healthcare settings.⁶ Kessler and McClellan (1996, 2002) and Dubay et al (1999, 2001) are relatively recent examples of papers that have found strong effects using US data, while others such as Beider and Hagen (2004) have failed to confirm such results. Overall, comprehensive surveys by Bovbjerg and Sloan (1998) and Mello and Brennan (2002) conclude that an empirical link has not been convincingly established. Of course, care may be oversupplied or undersupplied and it is possible to interpret an oversupply as ‘defensive medicine’; several authors have tried to identify this. Again, the results are mixed. Localio et al (1993) found statistically significant links between several obstetric procedures and previous claims experience (in New York State) but Baldwin et al (1995) could not confirm this result. Kessler and McClellan (1996) find a negative relationship between tort reform (i.e. reductions in liability) and costs of the compensation system: they interpret this as evidence of

and Peters, Jr (2008). Sloan and Chepke (2008) also provide an interesting discussion of these debates and their prospects.

⁶ Related empirical work is available in other settings that use tort liability as well. For example, automobile accidents are influenced by driver care and many jurisdictions operate tort-based compensation arrangements: see McKewin (1989), Devlin (1992) and Cummins et al (2001).

defensive medicine.⁷ Subsequent work by Sloan et al (1997) fails to find substantial evidence of any link between clinicians' previous claims history and (i) the choice between Cesarian and vaginal birth; (ii) the use of antenatal testing, and (iii) mothers' satisfaction with their treatment levels. In general, they find no significant differences as claims history changes, suggesting that liability risk (as measured by previous history) does not affect (and, in particular, raise) care levels.

In this paper, we examine the extent to which risk sharing arrangements in an enterprise liability context affect the quality of care supplied by hospitals. We use data from the National Health Service in England and Wales, but the issues of patient safety that we address are clearly common to other health systems. In order to capture the variation in liability risk across hospitals, we make use of explicit incentive mechanisms introduced as part of their risk sharing arrangements.⁸ The explicit incentives we study come in two forms: liability insurance deductibles that seek to share risk between each hospital and its insurer (the NHS Litigation Authority) and, then, premium discounts for attaining objectively set (and monitored) risk management standards – again offered by the NHSLA. Both deductibles and attained risk management standards differ across hospitals (and can change over time) according to decisions the hospitals themselves take and we have data on every hospital's arrangements in this respect over a five year period (2000-2004). In addition, we have uniquely matched these to hospital level data on hospital characteristics including measures of patient care. To be specific, we examine the association between hospitals' liability insurance deductible levels, the risk management standards they achieve and the utilization of diagnostic imaging tests – one type of diagnostic activity that we suggest can be helpful in reducing medical error and improving quality of care and patient safety.⁹

Fenn et al (2007) explored the relationship between the incidence of diagnostic tests and deductibles in a given year, but here we extend that analysis using a panel dataset

⁷ It is possible, as Danzon (2000b) explains, to attribute this result to the presence of managed care in the US. Thus, in a follow-up study, Kessler and McClellan (2000) control for managed care and find a reduced but still significant 'defensive effect'.

⁸ This distinguishes our paper from many other studies, where differential liability risk is proxied by differences in tort rules across States or over time, or by past claims histories. In addition, our paper examines the issue of liability risk and incentives for care in the context of a system-wide enterprise liability system, whereas other studies have focused more specifically on clinician-based insurance.

⁹ Such diagnostic tests have also been used by other authors as proxies for care: see, for example, Kessler and McClellan (1996). We discuss the role of diagnostic tests in a little more detail in Section 5.

covering the above five year period. At the beginning of this period deductibles varied across hospitals, but from 2002 onwards they became zero for all hospitals, after a policy decision to transfer all liability risk from hospitals to the NHSLA. By looking over a number of years during which these risk sharing arrangements were changing, we are able to observe variations of deductibles between hospitals prior to 2002, and a sudden shift over time in the level of excesses for all hospitals, after which the NHSLA increased its reliance on risk management standards as a means of influencing hospital behaviour. In effect, we have a “natural experiment” resulting from this policy change, with the added advantage that the experiment is embedded in a cross-sectional and time-series data set, raising the possibility of moving beyond a straightforward before/after study design and allowing us to explore the potential for using panel data methods.¹⁰

Our analysis focuses on the health care sector, where the role of tort and other mechanisms for encouraging good standards of care have been extensively discussed in recent years (eg Danzon 2000a,b; Department of Health 2003), but the methods and results may have implications for other public services. For example, local authorities (whether they be federal, state or county level) commonly face enterprise liability and insure themselves against liability claims, giving rise to similar issues to those that we study: can tort encourage care levels in the face of risk pooling arrangements? The public corporation arrangements found in many countries (for example in the rail or energy sectors) also raise similar institutional issues.

The paper is structured as follows. The next section provides some institutional background before the third explains our estimation methodology and a fourth presents our data. The fifth section contains results and a discussion of these before conclusions are drawn in Section 6.

2. The institutional background

In 1990 the NHS in England and Wales adopted responsibility for any mistakes of directly employed staff, the resulting “enterprise liability” system focusing claims onto hospital management rather than individual clinicians. Following this, in November 1995 the NHS Litigation Authority (NHSLA), a Special Health Authority, was established and

¹⁰ As is common in such cases, the quality of available instruments is central to such analyses; a point we address in Section 3.

became the main mechanism allowing NHS bodies to pool the costs of liabilities to patients. The NHSLA administers a number of different schemes, the two main ones being the Clinical Negligence Scheme for Trusts (CNST) and the Existing Liabilities Scheme (ELS). The former is a voluntary scheme to which all NHS hospital Trusts and primary care trusts (PCTs) currently belong (although membership is not compulsory), and covers clinical incidents occurring on or after the date when the Trust joined the scheme. Until April 2002, Trusts were able to select a deductible, and were responsible for handling and funding claims below that level. In April 2002 however, these lower value claims were “called-in”, and all CNST claims are now handled centrally by the NHSLA, although the Trust remains the legal defendant. The scheme is funded by ‘pay-as-you-go’ contributions from hospitals which are determined primarily on the basis of the risk classification of each hospital, which in turn is determined by the NHSLA using a formula based on numbers of clinical staff in different specialties to estimate expected NHS CNST claim costs.

The “call-in” of all claims in 2002 and the ending of risk sharing via excesses has potentially diluted hospital incentives to invest in patient safety (see Fenn et al, 2004). However, the NHSLA has also developed the use of experience- and merit-rating. The former has played a relatively small role in the assessment of contributions to date, in part because the scheme has not reached maturity and full claims history is still being acquired. In the case of merit-rating, each hospital is assessed against three levels of risk management standards every two years. Trusts which are assessed as complying with the standards are then given a discount on their scheme contribution for the following two financial years. The discounts on CNST contributions are 10% (level 1 compliance), 20% (level 2 compliance) and 30% (level 3 compliance). The average CNST contribution paid by NHS hospital Trusts in 2005/6 was approximately £1.7m, with many larger Trusts contributing significantly more than this. Achieving higher level risk management standards can therefore produce significant financial savings, potentially creating incentives for Trusts to invest in patient safety. The process by which risk management standards are assessed involves the review of a large number of processes across a wide spectrum of activities which might directly or indirectly influence diagnostic care. For example, current (2009-10) risk management standards for hospitals specify that organizations can demonstrate processes to ensure staff are adequately trained to use diagnostic equipment safely. Although they impose no requirements concerning the

actual use of clinical diagnostic tests, this has been identified a risk area for potential inclusion in future standards.

3. Estimation

As noted above, our intention is to explore the extent to which liability risk-sharing measures across NHS hospitals are associated with activities relevant to patient safety in these hospitals, such as the use of diagnostic measures. In particular, we intend to test whether hospitals choosing larger deductibles or higher risk management standards will choose higher diagnostic care levels, *ceteris paribus*, as these incentive mechanisms intend.

The data acquired for this project – described in the following section – constitute a panel of cross-sections (hospitals) over time (annual observations), and are in the form of counts of the numbers of diagnostic tests observed in each hospital in a given year. For each such measure, we therefore seek to estimate the relationship between these counts and the hospital's CNST excess level and/or risk management standard, controlling for throughput, casemix and caselength variations across hospitals. Moreover, because the hospital's choice of excess (prior to 2002), and its decisions with respect to risk management expenditure, are potentially choice variables, we also allow for regressor endogeneity using a dynamic GMM system estimator. We now outline this estimation strategy in more detail.

For a given hospital i , the process by which observed data are generated on the numbers of diagnostic tests in year t could be characterised as a Poisson process with a rate of occurrence, π_{it} . The observed total number of tests would clearly depend on a measure of patients' diagnostic requirements – that is, the number of treatment episodes for the hospital in a given year (T_{it}). Consequently, the expected number of tests in a given year at a given hospital would be $T_{it}\pi_{it}$ where π_{it} represents the mean probability of a test during a given treatment episode. Given an assumed Poisson process, this would imply that the observed number of tests (x_{it}) in hospital i in year t is distributed with density

$$f(x_{it}, T_{it}, \pi_{it}) = \frac{e^{-T_{it}\pi_{it}} (T_{it}\pi_{it})^{x_{it}}}{x_{it}!} \quad (1)$$

The parameter π_{it} can be modelled as a function of observed covariates and unobserved random variables. With a conventional loglinear specification of this function, we have

$$\pi_{it} = \exp(\beta_1 \mathbf{p}_{it} + \beta_2 d_{it} + \beta_3 e_{it} + \varepsilon_{it}) \quad (2)$$

Where $\mathbf{p}_{it}$ is a vector of measures capturing the hospital's location, casemix and activity rates (length of stay, bed utilisation rate). The variable d_{it} measures the risk management standard achieved by the hospital and variable e_{it} measures the deductible; β_1 , β_2 and β_3 are the associated coefficients. The error term ε_{it} measures the impact of unobserved heterogeneity in the underlying risk across hospitals: given the difficulty in capturing through observables all variations in case complexity, management efficiency, and other local factors, this is likely to be a significant determinant of inter-hospital variation in infection rates.

The availability of a panel of data on hospitals over a period of six years allows us to obtain valuable additional observations on diagnostic tests, increasing the power of the cross-sectional analysis as well as exploring changes in diagnostic care over time. In addition, however, the extra data also raises the possibility of addressing the unobserved heterogeneity across hospitals through specifying $\varepsilon_{it} = u_i + v_{it}$ and then taking into account these hospital-level effects in the estimation process

If we maintain the loglinear assumption for the relationship between the incidence of tests and observed covariates (equation (2)), a more general specification of the data generation process would be

$$\frac{x_{it}}{T_{it}} = \exp(\beta_1 \mathbf{p}_{it} + \beta_2 d_{it} + \beta_3 e_{it} + u_i + v_{it}) \quad (3)$$

Furthermore, if x_{it} is strictly positive, equation (3) can be transformed for estimation as

$$\ln\left(\frac{x_{it}}{T_{it}}\right) = \beta_1 \mathbf{p}_{it} + \beta_2 d_{it} + \beta_3 e_{it} + u_i + v_{it} \quad (4)$$

Assuming exogeneity of both ρ_{it} , d_{it} and e_{it} this log-linear model (with dependent variable the log of the rate of testing¹¹) could be consistently estimated using either a random-effect GLS estimator or an OLS fixed-effect estimator.

However, as pointed out earlier, we believe that d and e are potentially endogenous regressors due to the discretion hospital management has when deciding what level of deductible is appropriate, as well as what expenditure is required on risk management activities. If we can further assume that $E[\ln(\varepsilon_{it})|z]$ is a constant independent of an instrumental variable vector z , then even where d and e are shown to be endogenous it is possible to estimate consistently the parameters of the log-linear model at (4) using an instrumental variable approach. Critically, there are relatively few clear external possibilities available to us for instrumental variable purposes, and it is natural then to look to lagged values for instruments. In addition, it is possible that there is a degree of persistence in the reported rate of testing within a given hospital. If this is believed to be a significant factor, then the appropriate estimating equation will have a lagged dependent variable as one of the regressors. In the case of the loglinear version of the data generating process, we would have:

$$\ln\left(\frac{x_{it}}{T_{it}}\right) = \alpha \ln\left(\frac{x_{it-1}}{T_{it-1}}\right) + \beta_1 \rho_{it} + \beta_2 d_{it} + \beta_3 e_{it} + u_i + v_{it} \quad (5)$$

where the parameter α measures the degree of persistence.

In estimating (5) we have an additional source of endogeneity bias through the presence of the lagged dependent variable (“dynamic panel bias”). Moreover, in this case the standard fixed-effect estimator is unable to solve this problem using higher lags of the dependent variable as instruments. The usual solution to this problem is to resort to the use of first differencing of the data, and then to estimate the resulting equation using GMM (Arellano and Bond, 1991). The (differenced) lagged dependent variable remains endogenous, but now it is possible to use the lagged levels of the dependent variable as instruments. However, in circumstances where there is stochastic variation in the dependent variable around a relatively stable trend, it has been shown that difference GMM lacks efficiency because in that case lagged levels convey little information about

¹¹ Note that there may be more than one test per treatment episode, so the left-hand side of equation (3) is in the form of a rate, not a probability (i.e. it could in practice exceed one for some hospitals and some types of test). Hence a log-linear specification (rather than a logit) is appropriate.

future changes (Blundell and Bond, 1998). Drawing on Arellano and Bover (1995), Blundell and Bond show that an alternative is to estimate (5) as an untransformed equation in levels, using the differences in the lagged dependent variable as instruments. These differenced instruments are exogenous providing that they are uncorrelated with the fixed effects u_i .¹² They can be added to the instruments in levels from the difference GMM estimator to form a potentially more efficient system GMM estimator (Blundell and Bond, 1998). We have used this system GMM estimator below for each of the diagnostic tests for which we have data, and in each case we have assumed endogeneity in respect of the hospital's deductible and risk management standards. External instruments used are the hospital's total number of new claims for clinical negligence, and a measure of its financial strength (operating surplus as a percentage of overall expenditure). These instruments are entered contemporaneously, and are chosen on the intuitive basis that the claims experience and financial performance of hospitals may well affect their choice of deductible and risk management standard respectively, but should not directly influence the number of diagnostic tests felt to be clinically desirable. Additional internal instruments used are the contemporaneous and lagged values of the exogenous regressors, as well as the second and higher lagged values of the endogenous variables (i.e. the lagged dependent variable, the deductible and risk management standards). We are aware that this approach to identification potentially leads to a number of superfluous instruments. In particular, the deductibles for the years after 2002 are all set to zero, and are therefore exogenous, so the use of lagged values as instruments is unnecessary for these observations.¹³ A similar issue arises with respect to the risk management standards, as some hospitals are not re-inspected every year. However, while our approach therefore has some cost in terms of efficiency, the estimates should remain consistent, providing that tests of overidentifying restrictions do not reject instrument exogeneity. For robustness we also estimated alternative models in which the deductibles and risk management standards were treated as exogenous (these are available from the authors on request, but they did not substantively affect our findings as set out in Section 5 below).

¹² It can be shown that, under certain assumptions, this requirement is equivalent to one about the initial conditions of the data generating process (Blundell and Bond, 1998).

¹³ We are grateful to an anonymous referee for this point.

4. Data

Information on the deductibles chosen by each Trust in each year between 1995 and 2001 was provided by the NHSLA, and Table 1 shows the number of Trusts selecting particular deductibles in each year.

Table 1 near here

Prior to 1997, a deductible level of £100,000 was compulsory. Most Trusts opted for lower levels after choice was introduced in subsequent years, and from 2002 all Trusts had a £0 deductible as a consequence of the policy shift described above. Consequently during our period of analysis (2000 to 2004) there were two years during which hospitals experienced various different deductible levels, and three subsequent years when the deductibles were effectively zero.

The NHSLA also provided us with information on risk management scores for each hospital, consisting of the 2-yearly or yearly assessments which provide the basis on which discounts for risk management performance are assessed, and Table 2 shows the overall risk management levels attained by year from 2000 to 2005.

Table 2 near here

It is evident that as the assessment of risk management standards has developed, the number of Trusts on the lowest risk management scores of 0 or 1 has declined significantly, while the number attaining higher standard levels has increased.¹⁴

In order to capture variations in care levels across hospitals, we focused on indicators of diagnostic activity, specifically in relation to the use of imaging tests; these can be viewed as proxies for the degree of diagnostic care exercised by clinical decision-takers in each hospital, as well as the diagnostic resources made available to clinicians by hospital management.

¹⁴ The reductions in the total number of hospitals between 2000 and 2002 are largely accounted for by hospital mergers and consolidation

Table 3 shows data on the mean number of different types of diagnostic imaging activities undertaken by Trusts over the period 2000-2004. Figure 1 shows how the numbers of key imaging tests have been distributed across hospitals during this period.

Figure 1 and Table 3 near here

The variability in the number of events shown in the Figure is striking, but of course is driven in part by variations in activity levels in the respective hospitals (i.e. the number of treatment episodes), as well as different casemix and other factors.

In attempting to discover the additional role played by liability risk-sharing arrangements, we clearly need to take into account and control for these other factors, and we have done so by using indicators of throughput, case-mix and caselength, derived from Hospital Episodes Statistics (Department of Health, 2006). Table 4 shows the mean values of these control variables for the years 2003 and 2004.

Table 4 near here

5. Results and discussion

Our analysis explores the relationship between hospital CNST excess levels, risk management standards and the utilization of diagnostic imaging tests over the five year period from 2000 to 2004, with diagnostic tests viewed as one proxy for level of care. In the middle of this period (in 2002), as described above, the NHSLA implemented a policy shift which removed the ability of hospitals to choose an excess level, in effect reducing the deductible to zero for all hospitals. The results of our dynamic system GMM regressions are presented in Table 5. In all cases the first order autocorrelation is significant, as to be expected. The second order autocorrelation is insignificant in all cases but one (X-rays) where it is marginally significant at the 5% level. In all cases the J-test for overidentifying restrictions does not reject the null of instrument independence.

Table 5 near here

The table suggests that by far the most significant determinant of all the diagnostic tests we study is previous use (i.e. the one-period lagged dependent variable). In all cases the coefficient is strongly significant and below one, suggesting a significant degree of

persistence in the use of these tests. In particular, the high volume tests such as X-rays, CT scans and non-specific ultrasound scans have coefficients close to one, indicating a very slow adjustment process; these tests are prescribed in a fairly routine manner, and the hospital's budgeted capacity is presumably determined in part by reference to the previous year's expenditure. The impact of the hospital's mean length of stay also has a generally positive and significant effect on the rate of testing, particularly so for MRI scans and obstetric ultrasound scans. Presumably the frequency of these tests is sensitive to the duration of the patient's period of observation and monitoring while in hospital. As expected, hospitals with higher proportions of general surgery make more significant use of CT scans, MRI scans and X-rays, while those specializing in gynaecology and obstetrics make statistically more use of obstetric scans. Perhaps it is not surprising that, having controlled for these hospital characteristics, others (such as specialty and location) do not on the whole seem to play a significant role in the use of diagnostic and imaging tests.

Most important from our perspective are the coefficients on 'Deductible and 'CNST levels 2/3' – i.e. our measures of liability risk based on insurance deductibles and attainment of high risk management standards (and associated low insurance premiums). If either of these provides incentives to supply care, we might expect significant positive coefficients here. In fact, all the coefficients here are insignificant, though many have the 'correct' sign. This suggests that our measures of the incentives for care under risk sharing arrangements are not major determinants of diagnostic and imaging activity. In particular, it suggests that the policy shift in 2002 to transfer liability risk from hospitals to the NHSLA, and the subsequent emphasis on risk management discounts, has not had any discernible effect on diagnostic testing behaviour in NHS hospitals.

This result suggests a number of possible explanations. One possibility is that tort liability can be a powerful incentive for care but that the incentives inherent in the NHSLA's arrangements with hospitals are suboptimal. Although hospitals chose their own excesses prior to 2002, these were typically below £25,000 in most cases and this may have provided weak incentives. Similarly the range of risk management standards and discounts on offer may have been too limited. If this is true, our results have implications for policy makers seeking to incentive care – and perhaps for the merits of using a monopoly insurer when competitive pressures may have provided a wider menu of risk sharing arrangements.

Another possibility is that our proxies for care are not the best ones for identifying the effects of enterprise liability risk. In particular, diagnostic tests are perhaps more responsive to ‘clinician-targeted’ initiatives when, consistent with the NHS’s enterprise liability system, the incentive mechanisms explored here may be more appropriately thought of as ‘hospital-targeted’. If this is correct, we might expect to find stronger effects when proxying care by hospital initiatives such as infection control. These are legitimate points but it should also be noted that diagnostic care has been used elsewhere in the literature as a measure of care and, in principle, is clearly an observable proxy for the underlying latent variable which is a clinician’s desired level of confidence in their diagnosis. One possible response to this is that much of the literature using diagnostic tests to proxy care hails from the US, where they have sometimes been interpreted as good measures of excess or supplier-induced care. However, overall diagnostic rates are very much lower in the UK,¹⁵ rendering that interpretation less likely. Other institutional differences could alter the interpretation of diagnostic activity as a proxy for care levels in the UK. For example, in the US, privately insured care decisions may be more a matter of clinician choice than in a relatively ‘cash-constrained’ NHS. This may then be an additional reason why we find a weak link between these and mechanisms designed to incentivise care. In fact, though such differences should be acknowledged, they need not undermine our interpretation of diagnostic imaging. For instance, even cash-constrained hospitals can decide to allocate extra resources to measures of ‘care’ that can reduce liability risk. In addition, there is evidence that NHS physicians can exercise discretion in the application of such measures (for example, Gillan et al, 2001, discuss the discretionary use of radiology in treatment of lower back pain).¹⁶

Interestingly, mention of hospital decision making raises another issue: how do hospitals relay the incentives they face back to individual clinicians and, therefore, influence their subsequent care decisions: i.e. the internal operation of enterprise liability in the NHS? Legal proponents predict beneficial effects from adopting enterprise liability: for example, cost savings from not having to cite multiple defendants in an action, and ‘accuracy’ gains from lifting the case away from the clinician’s emotional and reputational involvement. Economic theory has also been used to examine the effects of enterprise

¹⁵ In 2007 MRI rates were 91 per 1,000 population in the US compared with 29 in the UK, and CT scans were 228 per 1,000 population in the US compared with 59 in the UK. (OECD 2009)

liability on care. Kornhauser (1982) provides the initial analysis to this effect: in a principal-agent setting with wealth constrained agents (clinicians), it may be optimal to switch liability to the deep pockets of the hospital, who offers a contract trading risk for care in return. Chu and Qian (1995) and Zeiler (2004) also consider the effects of liability rules in a principal-agent setting while Arlen and MacLeod (2005) provide a sophisticated model in which clinician expertise and MCO authority co-exist. (See also Arlen and MacLeod, 2003). They show that the presence of expertise encourages the MCO to review clinician activity and that the resulting distortions in care make it liable to transfer liability to the MCO. The burden of these models is that hospitals can solve the principal-agent problems they face and pass on to their clinicians the incentives to secure patient care that arise from liability rules.

In fact, it may not be straightforward in practice to address the principal-agent issues at the heart of the foregoing literature. As Williamson and others have pointed out, complex organizations can suffer from considerable inefficiencies, not least in terms of communications and internal incentives. Harris (1976) addresses such points explicitly to hospitals. He argues that the hospital is “two firms in one”: a medical staff and an administration, each with “its own objectives, decision variables and constraints” and, therefore, “locked in a noncooperative oligopoly-type game”. The resulting coordination problem can lead to suboptimal decision making and weak communication lines. Harris does not consider the implications of the analysis for the effects of enterprise liability on care, but a reasonable inference may be that hospitals are not good at passing the tort incentives to their employees and, as such, the variations in liability may have a weaker effect on care than expected.¹⁷

Thus, another potential interpretation of our results is that enterprise liability (at least, as operated in the NHS) may dilute the power of the incentives for care contained in insurance deductibles and risk management standards with premium discounts: hospitals may be too complex, or too resource constrained, to translate hospital-wide incentives into clinician-level ones. If this is the case, it suggests that enterprise liability rules need to be implemented in the context of careful organisational design.

¹⁷ Shortell et al (1976) provide some early empirical analysis on the role of organisational mechanisms in hospital performance (though not in direct relation to malpractice). Their results suggest that better specification of procedures and increased visibility within the hospital improves performance (costs and quality – measured by variables such as mortality rates and probabilities of complications).

6. Conclusions

The safety of patients is an important responsibility of health care providers, and significant compensation costs may arise if providers are negligent. An increasing trend is for liability for such compensation to be placed with the hospital rather than the individual clinician, a system known as ‘enterprise liability’. Hospitals in England and Wales have been subject to this system since 1990, and risk pooling arrangements have emerged subsequently allowing hospitals to transfer their liability risk to an agency - the NHS Litigation Authority. In this paper we have explored for the first time some of the mechanisms used by this agency to provide hospital management with financial incentives to take care. We have estimated the influence of these arrangements on the use of diagnostic imaging tests within hospitals, using a panel data set covering the period 2000-2004, during which period a policy shift took place leading to a form of “natural experiment”.¹⁸

As always in such analyses, much turns on the selection of suitable instruments and on the interpretation of key variables. We have acknowledged these points, and commented on our use of diagnostic imaging tests as a reasonable discretionary measure of care that may affect liability risk. Bearing both points in mind, our results suggest that the use of diagnostic tests did not respond to the incentives created during this period. We have speculated that certain types of patient care activity, including the use of diagnostic tests, may be less responsive to incentives placed at the level of the hospital by comparison with incentives placed at the level of the clinician. Nevertheless, to the extent that diagnostic care is constrained by management-provided resources, or ability to design/implement internal incentive mechanisms, these results may cast some doubt on the sensitivity of hospital management to financial incentives arising from their liability to patients. Our findings have implications for jurisdictions contemplating a move to enterprise liability, as well as wider implications for public sector organizations faced with financial incentives to improve service quality. As well as the need for continued research into the effects of incentive mechanisms to enhance care in health (and elsewhere), we believe that our results also suggest a research agenda relating to organisational design in health care. This would examine the role of such design in

¹⁸ In principle, the data allow an improvement on Fenn et al (2007) because the dynamic panel offers the chance to explore the use of diagnostic care across hospitals and over time.

DRAFT – DO NOT QUOTE

communicating and implementing high-level incentives within the organization, as required by effective enterprise liability arrangements.

References

- Abraham, K. and Weiler, P. (1994) "Enterprise medical liability and the evolution of the American health care system." *Harvard Law Review*, 108(2), 381-436.
- Arellano, M. and Bond, S. (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, 58(2), 277-297.
- Arellano, M. and Bover, O. (1995) "Another look at the instrumental variable estimation of error-components models." *Journal of Econometrics*, 68(1), 29-51.
- Arlen, J. and MacLeod, W. (2003) "Malpractice liability for physicians and managed Care Organisations." *New York University Law Review*, 78(6), 1929-2006.
- Arlen, J. and MacLeod, W. (2005) "Torts, expertise, and authority: Liability of physicians and Managed Care Organizations." *RAND Journal of Economics*, 36(3), 494-519.
- Baker, T. (2005) *The medical malpractice myth*. U of Chicago Press, Chicago.
- Baldwin, L., Hart, G., Lloyd, M., Fordyce, M. and Rosenblatt, R. (1995) "Defensive medicine and obstetrics." *Journal of the American Medical Association*, 274(20), 1606-1610.
- Beider, P. and Hagen, S. (2004) "Limiting tort liability for medical malpractice." Congressional Budget Office, Washington DC.
- Blundell, R. and Bond, S. (1998) "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, 87(1), 115-143.
- Bovbjerg, R. and Berenson, R. (2006) "Enterprise liability in the Twenty-First Century." In: Sage, W. and Kersh, R. (Eds), *Medical malpractice and the US health care system*, Cambridge UP, Cambridge UK.

Bovbjerg, R. and Sloan, F. (1998) "No fault for medical injury: Theory and evidence." *University of Cincinnati Law Review*, 67.

Chu, C. and Qian, Y. (1995) "Vicarious liability under a negligence rule." *International Review of Law and Economics*, 15, 305-322.

Cummins, J., Weiss, M and Phillips, R. (2001) "The incentive effects of no-fault automobile insurance." *Journal of Law and Economics*, 44(2), 427-464.

Danzon, P.M., 2000a. "Liability reform: Traditional and radical alternatives." In: Feldman, R.D. (Ed.), *American Health Care: Government, Market Processes and the Public Interest*. The Independent Institute, California.

Danzon, P.M., 2000b. "Liability for medical malpractice." In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*, 1B. Elsevier North Holland, Amsterdam.

Department of Health (2003): "Making Amends: A consultation paper setting out proposals for reforming the approach to clinical negligence in the NHS. A report by the Chief Medical Officer", Department of Health: London

Department of Health (2006 et seq): "Hospital Episode Statistics", Department of Health: London.

Devlin, R. (1992) "Liability versus no-fault automobile insurance regimes: An analysis of the experience in Quebec." In: Dionne, G. (Ed) *Contributions to Insurance Economics*, Kluwer Academic Publishers, Norwell MA.

Dubay, L., Laestner, R., Waidmann, T. (1999) "The impact of malpractice fears on cesarian section rates." *Journal of Health Economics* 18 (4), 491-522.

Dubay, L., Laestner, R., Waidmann, T. (2001) "Medical malpractice liability and its effects on prenatal care utilization and infant health." *Journal of Health Economics* 20 (4), 591-611.

Fenn, P.; Rickman, N.; Gray, A. (2004), "The economics of clinical negligence reform in England." *Economic Journal*, Vol.114 (496), pp. F272-292

Fenn, P.; Rickman, N.; Gray, A. (2007) "Liability, insurance and medical practice." *Journal of Health Economics* 26(5), 1057-1070.

Gillan, M. G. C.; Gilbert, F. J.; Andrew, J. E.; Grant, A. M.; Wardlaw, D.; Valentine, N. W.; Gregori, A. C. (2001) "Influence of imaging on clinical decision making in the treatment of lower back pain." *Radiology* 220, 393-399

Harris, J. (1976) "The internal organization of hospitals: Some economic implications." *Bell Journal of Economics*, 8(2), 467-482.

Kessler, D. and McClellan, M., (1996) "Do doctors practice defensive medicine?" *Quarterly Journal of Economics* 111 (2), 353-390.

Kolstad, C. Ulen, T. & Johnson, G. (1990): "Ex post liability for harm vs ex ante safety regulation: Substitutes or complements?" *American Economic Review*, 80(4), 888-901.

Kornhauser, L. (1982) "An economic analysis of the choice between enterprise and personal liability for accidents." *California Law Review*, 70(6), 1345-1392.

Leone Jr, A. (2006) "The virtual reality of enterprise liability." Britcher, Leone & Roth, LLC.

Localio, A., Lawthers, A., Bengtson, J., Hebert, L., Weaver, S., Brennan, T and Landis, J. (1993) "Relationship between malpractice claims and Cesarean delivery." *Journal of the American Medical Association*, 269(3): 366-373.

McEwin, I. (1989) "No-fault and road accidents: Some Australasian evidence." *International Review of Law and Economics*, 9: 13-24.

Mello, M. and Brennan, T. (2002) "Deterrence of medical errors: Theory and evidence for malpractice reform." *Texas Law Review*, 80.

Organisation for Economic Cooperation and Development (2009). "OECD Health Data 2009. OECD, Paris.

Peters Jr, P. (2008) "Resuscitating hospital enterprise liability." *Missouri Law Review*, 73, 369-397.

Shavell, S. (1984) "A model of the optimal use of liability and safety regulation." *RAND Journal of Economics*, 15(2), 271-280.

Shavell, S. (1987) *Economic analysis of accident law*. Harvard UP, Cambridge MA.

Shortell, S., Becker, S. and Neuhauser, D. (1976) "The effects of management practices on hospital efficiency and quality of care." In: Shortell, S. and Brown, M. (Eds.) *Organizational research in hospitals*, Inquiry monograph, Blue Cross Association, September, 90-107.

Sloan, F. and Chepke, L. (2008) *Medical malpractice*. MIT Press, Cambridge MA.

Sloan, F., Entman, S., Reilly, B., Glass, C., Hickson, G. and Zhang, H. (1997) "Tort liability and obstetricians' care levels." *International Review of Law and Economics*, 17, 245-260.

Steves, M. (1975) "A proposal to improve the cost to benefit relationships in the medical professional liability insurance system." *Duke Law Journal*.

Wallis, H., and Mulcahy, L. (2001) "Cause for Complaint: An evaluation of the effectiveness of the NHS complaints procedure." London, Public Law Project.

Weiler, P. (1991) "Medical Malpractice on Trial." Harvard UP, Cambridge MA.

Zeiler, K. (2004) "Medical malpractice and contract disclosure: An equilibrium model of the effects of legal rules on behaviour in health care markets." *Georgetown Law and Economics Research Paper* no. 539224, Georgetown University.

Tables and figures

Table 1: Number of Trusts by deductible level and year, 1995-2005

Year	Deductible Levels							Total
	£0	£10k	£25k	£50k	£100k	£250k	£500k	
1995	0	0	0	0	366	0	0	366
1996	0	0	0	0	383	0	0	383
1997	0	58	218	82	35	2	0	395
1998	0	123	179	67	27	1	0	397
1999	0	139	159	59	31	0	2	390
2000	0	133	131	47	26	0	1	338
2001	0	130	102	38	25	0	1	296
2002	245	0	0	0	0	0	0	245
2003	239	0	0	0	0	0	0	239
2004	239	0	0	0	0	0	0	239
2005	173	0	0	0	0	0	0	173

Note: figures in bold relate to the period analysed in this paper

Table 2: Frequency distribution of overall risk management levels, 2000-2005

Year	Risk management scores:				Total
	0	1	2	3	
2000	81	207	33	1	322
2001	44	208	38	1	291
2002	47	159	36	3	245
2003	12	176	46	5	239
2004	0	176	53	10	239
2005	0	90	73	10	173

Note: figures in **bold** relate to the period analysed in this paper

Table 3: Imaging activity by year: (Average number per Trust)

Year	CT scans	MRI scans	Obstetric scans	Non-obstetric scans	X-rays	Radio-isotope scans	Fluoro- scopies
2000	7,754	3,636	10,438	15,594	87,723	3,434	6,057
2001	8,881	4,127	11,376	17,745	95,218	3,467	6,330
2002	10,488	4,761	13,150	20,783	111,803	3,676	7,647
2003	11,228	4,834	11,559	21,601	111,454	3,232	6,879
2004	12,207	5,389	11,833	22,212	111,270	3,151	6,780

Table 4: Activity levels and casemix indicators, NHS Trusts 2003-2004

	2003	2004
Mean no. of finished consultant episodes	56,279	57,760
Mean length of stay (days)	12.3	12.6
Proportion of bed-days in:		
General medicine	0.169	0.172
General surgery	0.072	0.072
Gynaecology	0.024	0.023
Obstetrics	0.026	0.026
Paediatrics	0.049	0.048
Orthopaedics	0.079	0.079
Urology	0.015	0.015
Other surgery	0.044	0.046
Other medicine	0.222	0.233
Mental illness	0.237	0.242

Table 5: Diagnostic imaging tests in relation to CNST deductibles and risk management standards; System GMM regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	CT scans	MRI scans	Obst scans	Other scans	X-rays	Radio- iso	Fluoros copies
Ln(test rate)[t-1]	0.872*** (15.42)	0.634*** (4.55)	0.468** (2.87)	0.970*** (28.78)	0.953*** (47.24)	0.910*** (13.01)	0.761*** (5.50)
Ln(length of stay)	0.358* (2.36)	0.992** (3.06)	1.248** (3.26)	0.0661 (0.99)	0.140* (2.04)	0.250 (1.42)	0.465 (1.61)
Ln(bed utilization rate)	-0.0497 (-0.32)	0.242 (0.61)	-0.346 (-1.27)	-0.739 (-0.01)	0.0481 (0.46)	-0.133 (-0.49)	-0.171 (-0.75)
Deductible	0.080 (0.27)	0.452 (0.66)	0.402 (0.42)	0.099 (0.41)	0.176 (0.74)	0.148 (0.52)	0.712 (1.81)
CNST level2/3	-0.0063 (-0.28)	0.0199 (0.27)	-0.0807 (-1.39)	0.00261 (0.12)	0.00679 (0.32)	-0.0071 (-0.19)	-0.0139 (-0.41)
Proportion general surgery	2.748* (2.52)	4.179*** (3.31)	6.586*** (3.35)	0.938 (1.61)	1.442** (2.96)	1.321 (1.26)	3.074 (1.57)
Proportion gynaecology	0.566 (1.08)	0.381 (0.28)	5.871** (2.60)	-0.0230 (-0.07)	-0.277 (-1.24)	-0.0204 (-0.04)	0.450 (1.07)
Proportion obstetrics	0.208 (0.35)	0.949 (0.57)	6.586** (3.14)	0.0248 (0.09)	-0.0496 (-0.24)	0.194 (0.34)	0.0872 (0.30)
Proportion paediatrics	0.336 (1.27)	1.562 (1.85)	1.144 (1.08)	0.0965 (0.43)	0.256 (1.27)	0.179 (0.86)	1.099 (1.60)
Proportion trauma & orthopaedics	-0.0815 (-0.44)	1.354 (1.65)	5.066** (2.84)	0.296 (1.94)	0.272* (2.04)	-0.0810 (-0.48)	0.636 (0.95)
Proportion urology	0.215 (0.20)	4.265 (1.82)	1.678 (0.60)	-0.108 (-0.15)	-0.0918 (-0.13)	0.457 (0.42)	3.036 (1.33)
Proportion other surgery	0.262 (1.37)	1.041 (1.51)	1.963* (2.39)	0.164 (0.96)	0.197 (1.58)	0.375 (1.29)	1.524 (1.42)
Proportion other medicine	0.176 (1.82)	0.254 (1.27)	0.515* (2.27)	-0.0267 (-0.33)	0.0372 (0.60)	0.129 (0.94)	0.141 (1.04)
Proportion psychiatry	-0.622** (-2.60)	-0.637 (-1.52)	0.600 (1.21)	-0.0350 (-0.13)	-0.234 (-1.06)	-0.156 (-0.59)	0.0337 (0.07)
Year=2002	0.0218 (0.89)	0.0607 (1.38)	0.103* (2.25)	-0.0128 (-0.50)	-0.0002 (-0.01)	0.0414 (1.23)	0.0395 (1.24)
Year=2003	0.0258 (0.99)	0.0912 (1.13)	0.140** (2.83)	-0.0219 (-0.81)	-0.0170 (-1.01)	-0.0319 (-1.04)	0.0626* (1.97)
Year=2004	0.0340 (0.89)	0.175 (1.88)	0.243** (2.80)	-0.0337 (-1.12)	-0.0197 (-0.64)	-0.0289 (-1.01)	0.0965* (2.41)
Single specialty hospital	0.233 (1.65)	0.598* (2.47)	-0.909 (-1.47)	0.0562 (1.09)	0.0778 (1.40)	0.133 (0.89)	0.0205 (0.16)
Teaching hospital	0.0365 (1.30)	0.100 (1.39)	0.0151 (0.21)	0.0142 (0.68)	0.0110 (0.67)	0.0122 (0.35)	0.0345 (0.80)
East of England	0.0879 (1.12)	0.130 (0.67)	0.196 (0.75)	0.00154 (0.02)	0.0555 (1.29)	-0.0560 (-0.85)	0.253 (1.51)
London	0.0544 (0.71)	0.0619 (0.31)	0.143 (0.57)	0.0354 (0.46)	0.0304 (0.69)	-0.0590 (-0.76)	0.129 (0.82)
North East	0.0898 (1.24)	0.135 (0.78)	0.114 (0.46)	0.0484 (0.67)	0.0586 (1.39)	-0.0525 (-0.75)	0.254 (1.40)
North West	0.0816 (1.12)	0.0763 (0.46)	0.143 (0.60)	0.0545 (0.92)	0.0517 (1.35)	-0.0285 (-0.40)	0.212 (1.29)
South Central	0.102	0.239	0.145	0.0517	0.0507	-0.0767	0.337

	(1.20)	(0.91)	(0.51)	(0.68)	(1.06)	(-0.80)	(1.69)
South East	0.0927	0.0326	0.0244	0.0364	0.0533	-0.0170	0.190
	(1.19)	(0.20)	(0.10)	(0.54)	(1.08)	(-0.21)	(1.26)
South West	0.142	0.245	0.184	0.0221	0.0561	-0.0457	0.307
	(1.82)	(1.15)	(0.80)	(0.32)	(1.29)	(-0.72)	(1.70)
West Midlands	0.126	0.0313	0.277	0.0382	0.0440	-0.0431	0.277
	(1.64)	(0.19)	(1.19)	(0.61)	(1.05)	(-0.61)	(1.70)
<i>N</i>	592	577	531	613	616	508	599
N of hospitals	166	165	150	169	172	140	168
χ^2	1320996	242931	202160	2593996	5979696	975181	775491
AR(1) test statistic	-3.576	-2.811	-1.245	-3.328	-3.053	-3.178	-1.467
p> AR(1)	0.00034	0.00494	0.213	0.00087	0.00226	0.00148	0.142
AR(2) test statistic	1.535	-1.549	1.422	0.926	1.998	-0.0074	1.124
p> AR(2)	0.125	0.121	0.155	0.355	0.0457	0.994	0.261
No. of instruments	68	67	67	67	67	67	68
Hansen J statistic	50.95	48.77	45.56	38.61	37.46	41.34	45.56
p>J	0.137	0.161	0.252	0.533	0.585	0.412	0.288

Robust *t* statistics in parentheses* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1: Selected diagnostic imaging rates as distributed across NHS hospital Trusts, 2000-2004