Centre for
Health Service Economics
& Organisation

Hospital Standardised Mortality Ratios (HSMRs) and all cause mortality

Barry McCormick | Jonathan White

Working Paper No. 1, September 2011

HSMRs and all cause mortality

Barry McCormick and Jonathan White

Working Paper No. 1

First published in September 2011 by:
Centre for Health Service Economics & Organisation (CHSEO)
5th Floor Zone B, Skipton House, 80 London Road, London SE1 6LH
Nuffield College, New Road, Oxford OX1 1NF

Corresponding author(s):

Barry McCormick
Director, Centre for Health Service Economics & Organisation
barry.mccormick@chseo.org.uk

Jonathan White
Project Lead, Centre for Health Service Economics & Organisation
jonathan.white@chseo.org.uk

Abstract

Objective: To assess whether the Imperial College/Dr. Foster Hospital Standardised Mortality Ratio (HSMR), which aims to indicate hospital safety, describes genuinely higher-than-expected death rates rather than confounding statistical issues.

Design: Analysis of the influence of hospital safety on all cause mortality in the PCTs they serve. If HSMRs measure safety, and are not wholly confounded by statistical issues, PCT all cause mortality would be expected to increase with local hospital HSMRs. Concern that HSMRs do not incorporate variation in 'place of death' between areas (resulting in a failure to fully case-mix adjust) is investigated by examining the relationship between PCT *out-of-hospital* mortality and local HSMRs.

Setting: England.

Population: Mortality in England 2005-2009. Census and other administrative data is used for the purpose of socio-economic controls.

Main Outcome Measures: The statistical significance and size of the influence of local HSMR rates in models of all-cause (i) total PCT mortality and (ii) outside-of-hospital PCT mortality.

Results: A statistically significant ($p<0.10$) positive association is found between HSMR rates and PCT all cause mortality, controlling for socio-economic influences. However, the coefficient estimates are only around one fifth of the size suggested by the view that the HSMR measures higher-than-expected deaths. A highly significant negative association is found between out-of-hospital mortality rates and local hospital HSMRs.

Conclusions: PCTs served by high HSMR hospitals have more deaths after controlling for socio-demographic factors. This suggests that HSMRs at least partly reflect either safety, as advocates maintain, or other unobserved case-mix factors, rather than being a statistical artefact explained by (for example) variation in coding practice. However the scale of additional all cause mortality is less than the HSMR values imply. This discrepancy may reflect the finding that variation in place of death influences the HSMR; out-of-hospital mortality is significantly lower in PCTs served by high HSMR hospitals. This is consistent with part of the estimated value of the HSMR reflecting a sampling bias from failing to control for large spatial variations in the fraction of near death patients admitted to hospitals. HSMRs remain a valid quality metric but further methodological development to address sampling variation appears worthwhile.

We are most grateful to Brian Jarman, Martin Hensher, David Mant and Mike Fleming for reviewing earlier manuscripts, and to Steven Middleton of Dr Foster Intelligence for providing us with the PCT-level HSMRs. The analysis was carried out using the Intercooled STATA 8.0 statistical package and Microsoft Excel 2002.

Table of contents

1. Introduction

Most policies to enhance the quality of care hinge upon the reliable measurement of clinical quality. As a consequence, the value of better information about healthcare quality is high. In this paper we consider one of the most well-known investigative tools for hospital quality to provide new evidence appertaining to both its central claim to measure above-expected deaths, and concerning sampling bias in comparing hospital safety.

The wide variation in hospital mortality rates has contributed to extensive discussion of how far mortality data can be analysed to reflect the quality of hospital care. Critical to this discussion are Hospital Standardised Mortality Ratios (HSMRs), which were introduced in an influential paper by Jarman et al.[1] and have proven controversial. The major criticisms made of HSMRs are fourfold. First, it has been suggested that variation in the HSMR is driven by variations in hospital coding practice. For example, a hospital that codes fewer co-morbidities would be allocated a higher (worse) HSMR because of this apparently less averse case-mix, even if its quality level and patient intake were actually identical to its peers[2]. Secondly, it has been argued that because some areas have less provision outside of hospital for the care of seriously ill patients (e.g. fewer hospices), hospitals in these areas would therefore have more inpatients at risk of death, leading to a higher mortality rate and a higher HSMR. Thirdly, it is argued that local patient populations may differ in non-measured ways and that hospitals serving more vulnerable local populations are likely to be measured with higher than expected mortality rates (the 'case mix adjustment fallacy'[3-4]). Fourthly, it has been argued that different HSMR methods give inconsistent rankings of quality of care when compared to each other or to other quality metrics[5-7].

Nonetheless, HSMRs are comparatively easy to compute from administrative data, have enabled hospitals to target critical quality improvement initiatives (with guidance available from internationally influential organisations such as the Institute for Healthcare Improvement), have highlighted clinical failings[8] and whilst imperfect may still be helpful when measuring hospital performance – particularly when used alongside other quality measurements. They are already used in several countries including the United States, Canada, the Netherlands, Sweden and England (with which this paper is primarily concerned, and where a new Summary Hospital-level Mortality Indicator has recently been agreed upon).

The key proposal of the HSMR model is that a measure of safety is provided by the scale of higher-than-expected in-hospital dying. This in turn requires a model which provides a forecast of deaths at each hospital based on hospital and patient characteristics, so that the difference between actual and forecast deaths may be calculated. (These differences reflect the 'error' terms in the hospital mortality equation.) However, if the HSMR is instead partly driven by variation between hospitals in coding practice or 'place of death' arrangements (neither of which are associated with higher overall mortality), this suggests that the HSMR measure may only create an illusion of higher-than-expected dying. These two criticisms raise the possibility that the forecasting model is no longer valid. Coding variation suggests that identical patients are being described differently. 'Place of death' variation suggests that the model is failing to control fully the differences in the patients admitted to different hospitals: for example, certain dying patients in some PCTs would be admitted to a hospice not a hospital. Since these patients are not a random sample of all admitted patients, controls for this variation are required if the samples are to be balanced.

This paper gives new evidence regarding whether high HSMRs genuinely reflect the scale of higher-than-expected dying that is currently suggested, and how far those using HSMRs to estimate the scale of unexpected dying should consider how in PCTs served by high HSMRs there may be offsetting and unexpectedly low out-of-hospital dying. It also discusses the direction in which the forecasting model in place requires further to development to remove sampling bias.

2. Methods

If hospital safety is, to some extent, being captured by the HSMR statistic then it provides a metric that should help to explain the rate of all cause mortality in the PCT being served by the hospital concerned. This is now explored.

2.1. All cause mortality and hospital safety: a statistical framework

We assume that all cause PCT mortality may be explained by an exponential function of a vector of socio-economic variables related to PCT *i,* $\underline{z}^i$, with parameters, $\underline{\beta}$ and an intercept, α. Consider also a "perfect" HSMR-like hospital safety measure, $\hat{h}^i$, where each percentage point increase implies that (in area *i*) hospital deaths are one percentage point higher than expected. This influence will also bear on all cause mortality since hospital mortality is an important component of all mortality, and leads to the following equation,

$$\text{ln (all cause mortality rate for area } i) \quad = \alpha + \gamma \ln(\hat{h}^i) + \underline{\beta}\underline{z}^i + \varepsilon^i ,$$

where ε^i is a normally distributed error term. Because both the all cause mortality rate and the hospital safety measure are both entered using natural logarithms, γ will represent the percentage change in all cause mortality for a one-percent worsening in the measure of hospital safety. Using empirical data, the parameters α, $\underline{\beta}$ and γ can be estimated through multiple regression analysis. Consider now the value of γ.

If all deaths occur in hospital, and all data are perfectly measured, γ could be expected to equal one. In other words, if the hospital safety metric were correct and suggested that a given hospital mortality rate was 10% higher than it should be, then the equation would predict that all cause area mortality would increase by the same percentage. However, because only a fraction of deaths occur in hospital, γ would instead equal that fraction, as reduced hospital safety can only affect the mortality of hospital patients. Current evidence suggests that about 50% of deaths occur in hospital [9] so that, in the absence of other considerations, a parameter estimate of 0.5 would be expected. However, two other influences on γ should be mentioned:

> (i) When mortality data covers a longer time period – in the study here it covers three years – some patients who pass away due to poor hospital safety will otherwise die within the time period anyway due to their poor initial health. This means that their death is recorded in the all cause PCT mortality data regardless of treatment safety, so excess hospital deaths due to unsafe

treatment will not always affect the measured all cause mortality rate. The larger this effect is, the smaller the estimated value of γ.

(ii) If the hospital safety measure is negatively associated with the rate of deaths outside of hospital, that will also reduce the value of γ. Such a negative relationship would not be driven by hospital safety, but could instead be driven by variation in the level of end-of-life provision available outside of hospital. If some areas have greater such provision and the HSMR does not fully adjust for this, they will have lower HSMRs (as hospitals are dealing with fewer seriously ill patients) at the same time as higher mortality outside of hospital, leading to a lower value of γ in national analyses.

The following analysis uses a PCT-level HSMR measure h^i (developed by the Dr. Foster Unit and kindly supplied by Dr. Foster Intelligence) as a proxy for the perfect hospital safety measure $\hat{h}^i$, and estimates the size and statistical significance of γ. It is acknowledged that these HSMR measures are not intended for use as a direct measure of hospital safety, but rather as an indicator of where hospital mortality might be higher than expected.

2.2. Sources of data used to estimate the regression framework

All age all cause mortality data is taken at PCT level from the NHS National Centre for Health Outcomes Development (NCHOD). Rates per 100,000 population have been obtained for 2005-2007, 2006-2008 and 2007-2009, both with and without direct standardisation for age and sex. The data are presented for periods of three calendar years to reduce the influence of random noise and are linked to PCT via individuals' place of residence.

HSMR data has been kindly supplied for this paper by Dr Foster Intelligence to match the time periods of the all cause mortality data (2005-2007, 2006-2008 and 2008-2009). The Dr. Foster Unit calculated these HSMRs by dividing the actual number of in-hospital deaths for residents of a given PCT by the number of in-hospital deaths predicted by a patient-level logistic regression model. Explanatory factors included sex, age (in 5-year bands), co-morbidities (based on Charlson score) and elective/non-elective status, plus a number of other factors.

Control factors that are associated with all age all cause mortality have been obtained from a variety of sources which are set out in Annex A. The factors include the percentage of the population in various socio-economic, education and ethnic categories, alongside population density, the crime rate, median earnings and the percentage of the population who are income deprived. Because this data is not always available at the PCT geography, some of the data is from similar geographies (e.g. Local Authorities or the pre-2006 PCT geography) from which population weighted averages are computed at the PCT level. In such cases, the 2007 GP attribution dataset and population data from the 2006 PCT reorganisation are used as weights. In addition to the main control variables, data on the fraction of the PCT population in 5-year age bands is obtained from NCHOD. This data is used to adjust for age when using unstandardised mortality rates.

The number of in-hospital deaths by PCT and gender is obtained for 2007-2009 from Hospital Episode Statistics. Unstandardised out-of-hospital mortality rates are computed from this by subtracting in-hospital deaths from the total number of deaths

set out in the NCHOD data and dividing by population. The resulting data shows 51% of deaths occurring in hospital, which closely matches the figure (49%) from unpublished Office for National Statistics data for 2008. This figure of around 50% of deaths occurring in hospital is helpful in interpreting regression coefficients.

The dataset is summarised in Annex B. The salience of the control variables is clear from this dataset; in a regression of directly age and sex standardised 2007-2009 all age all cause mortality on the main control variables, the R-squared value is 85%, implying that they explain a large part of the variation in all age all cause mortality. The wide variation in all cause mortality is also clear from the dataset, as is notable variation in the HSMRs.

3. Results

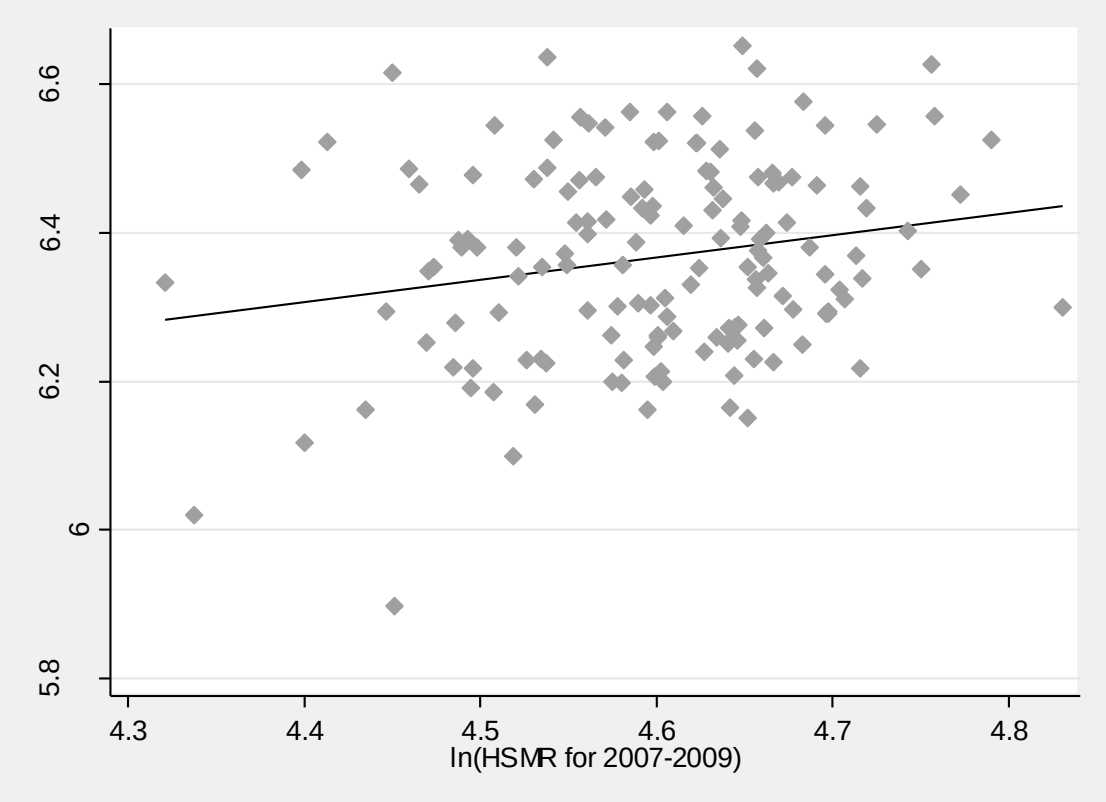
Using the framework set out above, bivariate and multivariate regression analysis are used to evaluate the existence and strength of a positive relationship between all cause mortality rates and HSMRs. As set out earlier, if such a relationship exists, HSMRs cannot solely be driven by variation in coding factors or place of death (as neither of these factors would affect overall area mortality). The coefficient on the HSMR is expected to be less than the fraction of deaths that occur in hospital.

3.1. Bivariate regression analysis

A simple bivariate association (Ordinary Least Squares) is first calculated between the 2007-2009 HSMR and directly age/sex standardised all age all cause mortality data, both in natural logarithms. A statistically significant positive relationship is found, as shown by the following regression equation and in Figure 1:

$$\ln(\text{PCT all cause mortality}) = 4.984219 + 0.30041 * \ln(\text{HSMR})$$

Figure 1: HSMR against standardised all age all cause mortality, both in natural logarithms (covering the three year period 2007-2009)



Whilst the coefficient on the HSMR is statistically significant at the 5% level (p-value 0.013), control variables have not yet been added. Nonetheless, the identified relationship is not dependent on a logarithmic specification; in a linear model the HSMR coefficient remains significant at the 5% level (p-value 0.026).

3.2. Main regression analysis (multivariate)

The bivariate regression from above is now repeated but with the inclusion of the PCT-level control variables, to examine whether the identified relationship remains. As noted earlier, these variables concern metrics for income deprivation and earnings, ethnicity, occupational class, educational level, and local social fragmentation, as captured by crime rates. They explain much of the variation in all cause mortality, with an R-squared of 85%. Ordinary Least Squares is used with robust standard errors, as implemented using the 'robust' switch in the Stata 8.0 statistical software. This method is more robust to issues such as heteroskedasticity.

Table 1 contains three main regressions – one where the all cause mortality and HSMR variables relate to 2005-2007, one for 2006-2008 and one for 2007-2009. Each main regression is repeated using backward stepwise regression (with an inclusion threshold of $p<0.2$) to identify the implications of dropping the less significant variables, giving six regressions in total.

Table 1: *Parameter estimates of models of directly age/sex standardised all age all cause PCT mortality rates.*

		Dependent variables - each column represents one regression					
Explanatory variables (below)		**ln(rateAllCause), standardised, 2005-2007**		**ln(rateAllCause), standardised, 2006-2008**		**ln(rateAllCause), standardised, 2007-2009**	
		Full	**Stepwise**	**Full**	**Stepwise**	**Full**	**Stepwise**
Intercept	Coefficient	5.853538	5.861324	5.844744	5.666497	5.594366	5.708616
	P-value	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
% income deprived	Coefficient	0.0186909	0.0197964	0.0203406	0.0205089	0.0207081	0.0217145
	P-value	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Population density	Coefficient	-0.0010053	-0.0012818	-0.0008743	-0.0012449	-0.0007197	(Dropped)
	P-value	0.160	0.055*	0.243	0.076*	0.307	
% Indian	Coefficient	-0.0023749	-0.003931	-0.0023989	(Dropped)	-0.0026603	-0.0037278
	P-value	0.351	0.000***	0.364		0.300	0.001***
% Pakistani and Bangladeshi	Coefficient	0.0015534	(Dropped)	0.001176	(Dropped)	0.0003635	(Dropped)
	P-value	0.525		0.650		0.892	
% Black Caribbean	Coefficient	-0.0104297	-0.0084372	-0.0081591	-0.0072438	-0.0064598	-0.010169
	P-value	0.492	0.037**	0.597	0.096*	0.716	0.034**
% Black African	Coefficient	-0.0025981	(Dropped)	-0.0028165	(Dropped)	-0.0028305	(Dropped)
	P-value	0.690		0.681		0.710	
% Chinese	Coefficient	-0.0187568	(Dropped)	-0.0260673	(Dropped)	-0.0247268	-0.0438146
	P-value	0.584		0.505		0.552	0.183
% Other Asian	Coefficient	-0.0102414	(Dropped)	-0.0137717	-0.0279602	-0.013988	(Dropped)
	P-value	0.541		0.436	0.001***	0.398	
% Other Black	Coefficient	0.0292696	(Dropped)	0.0150175	(Dropped)	-0.0022452	(Dropped)
	P-value	0.783		0.890		0.986	
% NS-SEC class 3	Coefficient	0.0251426	0.0234304	0.0254145	0.0266427	0.0269803	0.0144578
	P-value	0.013**	0.000***	0.015**	0.000***	0.011**	0.018**
% NS-SEC class 4	Coefficient	-0.0089806	-0.0086207	-0.0091686	-0.0093685	-0.0058426	-0.0146856
	P-value	0.206	0.104	0.218	0.080*	0.412	0.008***
% NS-SEC class 5	Coefficient	0.0269971	0.0257336	0.0302674	0.027279	0.0284697	0.0154808
	P-value	0.042**	0.006***	0.024**	0.004***	0.038**	0.125
% NS-SEC class 6 and 7	Coefficient	-0.0010822	(Dropped)	-0.0017492	(Dropped)	0.0016884	(Dropped)
	P-value	0.876		0.806		0.825	
% NS-SEC class 8, never worked	Coefficient	-0.0250673	-0.0254849	-0.0272515	-0.021934	-0.0230726	-0.0289436
	P-value	0.052*	0.003***	0.051*	0.027**	0.101	0.001***
% NS-SEC class 8, long term unemployed	Coefficient	-0.0096429	(Dropped)	-0.0169089	(Dropped)	-0.0236674	-0.0409081
	P-value	0.717		0.528		0.407	0.077*
% NS-SEC unclassified, students	Coefficient	0.0229923	0.0187814	0.0251013	0.0167437	0.0251107	0.0278094
	P-value	0.011**	0.006***	0.007***	0.022**	0.008***	0.001***
% NS-SEC unclassified, other	Coefficient	0.0020012	(Dropped)	0.0014905	(Dropped)	0.0025021	(Dropped)
	P-value	0.697		0.785		0.650	
Median earnings	Coefficient	-0.0004392	-0.000413	-0.0004149	-0.0004171	-0.0004471	-0.0004422
	P-value	0.031**	0.009***	0.054*	0.018**	0.030**	0.011**
% highest qualification 5 GCSEs	Coefficient	-0.0095695	-0.010285	-0.0097685	-0.0105432	-0.0080431	(Dropped)
	P-value	0.067*	0.022**	0.070*	0.023**	0.146	
% highest qualification 2 A-levels	Coefficient	-0.0340222	-0.0288762	-0.0362585	-0.0231247	-0.0338471	-0.0378701
	P-value	0.008***	0.001***	0.006***	0.018**	0.013**	0.001***
% highest qualification first degree	Coefficient	0.0055992	0.0046785	0.0055539	0.0062144	0.0073938	0.0054988
	P-value	0.143	0.050*	0.164	0.005***	0.075*	0.027**
% highest qualification other	Coefficient	-0.0170565	-0.0166652	-0.0158872	(Dropped)	-0.0100938	(Dropped)
	P-value	0.175	0.111	0.211		0.433	
Crime rate	Coefficient	0.0008943	0.0006934	0.0007905	0.0006999	0.0009549	0.0008927
	P-value	0.102	0.155	0.193	0.195	0.118	0.073*
HSMR	Coefficient	0.0974394	0.1010456	0.0913394	0.0893985	0.0858502	0.1019662
	P-value	0.047**	0.056*	0.101	0.115	0.118	0.053*

*The p-value is provided below each coefficient. *** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level. NS-SEC denotes the National Statistics Socio-Economic Classification.*

The Table 1 results show that the coefficient of interest (on the HSMR) is significant at the 5% level in one of the regression equations, significant at the 10% level in two more, and close to significance at the 10% level in the remaining three equations. In

two of the three time periods, the HSMR coefficients become more statistically significant when stepwise regression is used, whereas in 2006-2008 the coefficient becomes very slightly less statistically significant.

Furthermore, the coefficient values are all close to 0.1. As predicted in the Framework section, this is less than 0.5 (the proportion of patients who pass away in hospital). Individual F tests that these six coefficients equal 0.5 are rejected at the 1% level. Possible reasons already stated for this include the 3-year period of the mortality data and a possible link with outside-of-hospital mortality rates, which could both contribute to a coefficient of less than 0.5. That the coefficients only equal 0.1 implies that only one fifth (0.1 divided by 0.5) of the extra deaths suggested by a higher HSMR can be identified in the all cause mortality rate. Further analysis of outside-of-hospital mortality rates is included later on in this paper to investigate whether a negative relationship between outside-of-hospital mortality rates and the HSMR can partly explain this difference.

The control variables largely follow expected signs with negative effects on mortality for the measures of income deprivation and higher earnings, with certain of the lower status social classes, and with less education than GCSE/O-levels, although graduates appear to have mortality rates similar to poorly educated members of the population.

A variable denoting the number of whole-time equivalent GPs per 100,000 population was originally included in the above regressions, but was dropped because of its lack of statistical significance. Otherwise, all other variables (where appropriate, groups of variables) are statistically significant or close to statistical significance. Separately, an earlier analysis used PCT-level HSMR rates derived through spending-weighted averaging of 2005-2007 Trust-level HSMR rates, rather than using PCT-level HSMR rates directly from Imperial College/Dr. Foster Intelligence. The results were similar to in Table 1; the HSMR coefficient was significant at the 5% level both with and without stepwise regression, with a value of around 0.13.

3.3. Diagnostic testing on the regression equations

Two statistical tests are applied to the results using inbuilt functions from Stata 8.0. The Regression Specification Error Test (RESET, a test of omitted variables set out by Ramsey[10]) indicates that all six of the main regressions (including their stepwise variants) are misspecified at the 5% level. The Shapiro-Wilk W test, a test of normality[11-12], is then applied to the residuals from each of the main regressions, as normality is a key assumption of Ordinary Least Squares regression. In two of the six regressions, the null hypothesis of normality can be rejected at the 10% level. However, results below show that the statistical test results improve markedly once two outlying observations are dropped, whilst coefficient sizes and statistical significance remain similar.

3.4. Sensitivity analysis: dropping influential outliers

Examination of residual-versus-fitted-value plots suggests that Kensington & Chelsea PCT and Westminster PCT are influential outliers with unusually low all age all cause mortality. In a first piece of sensitivity analysis, the regressions are repeated without these observations to test the impact on the results. The outcome of the statistical tests improves markedly, with RESET tests that reject misspecification at the 10% level and higher and Shapiro-Wilk tests that reject normality of residuals at the 47%

level and higher. The statistical significance of the coefficients increases in five out of six cases whilst coefficient sizes remain close to 0.1. The results are summarised in Table 2.

3.5. Sensitivity analysis: gender-specific multivariate analysis of HSMR and unstandardised all age all cause mortality

As a second piece of sensitivity analysis, instead of using directly age and sex standardised all age all cause mortality (as in the above regressions), separate results tables were produced for men and women where the age standardisation is performed within the regression itself. This is conducted using an unstandardised death rate as the dependent variable, the previous set of controls, and additional explanatory variables capturing the fraction of the relevant gender under 1, the fraction aged 1-4, then the fraction in each 5 year age band up to the age of 85. The over-85s are the omitted group for each gender. As well as testing the sensitivity of the results to the choice of standardisation technique, this also enables comparison of the coefficient sizes between the male-specific and female-specific regressions. As before, stepwise estimation is also included, but with a restriction that the age variables (which are included for standardisation) are not dropped.

The male 2005-2007 results are statistically significant at the 10% level and become statistically significant at the 5% level once the aforementioned outliers are dropped. By contrast, the female results are not statistically significant in any of these permutations. Statistical significance at the 10% level is not observed for the male or female results in 2006-2008 or 2007-2009. The coefficients are summarised in Table 2.

The results fail to pass the RESET test even after the outliers identified in the main regression have been dropped. As is sometimes the case with the RESET test, there is no obvious resolution to the problem in this particular context. Additionally, the male results for 2005-2007 and 2006-2008 are not found to be normally distributed using a Shapiro-Wilk W test at the 10% level, although this is not the case once the aforementioned outliers have been dropped.

3.6. Summary of the results

Table 2 presents the coefficients of interest from the results so far. Whilst the results are not significant at the 5% (or even 10%) level in every case, most of the less significant coefficients are concentrated amongst the gender-specific internally standardised regressions. Despite this, the coefficients of interest are consistently positive. As discussed earlier, the findings provide support to the view that HSMR variation reflects a component of higher-than-expected dying and is not entirely driven by variation in coding practices and place of death.

The coefficients are also all close to 0.1, and notably less than the proportion of patients who pass away in hospital (0.5). Only around one fifth of the extra deaths detected by the HSMR can therefore be detected in the all cause mortality rate, which may be due to the 3-year period of the mortality data, and a possible link with outside-of-hospital mortality rates.

Table 2: Summary of parameter estimates of the influence of the HSMR in sensitivity modelling of PCT mortality rates

				HSMR coefficient p-value			
Gender	**Age/sex standardisation**	**Stepwise**	**Outliers removed**	**2005-2007**	**2006-2008**	**2007-2009**	**Full results**
Both	Direct (external)	No	No	0.0974394 0.047**	0.0913394 0.101	0.0858502 0.118	Table 1
Both	Direct (external)	Yes	No	0.1010456 0.056*	0.0893985 0.115	0.1019662 0.053*	Table 1
Both	Direct (external)	No	Yes	0.1132211 0.012**	0.0977197 0.044**	0.0912878 0.070*	On request
Both	Direct (external)	Yes	Yes	0.0992444 0.039**	0.0831866 0.082*	0.0791358 0.083*	On request
M only	Internal	No	No	0.1151612 0.056*	0.0767988 0.189	0.0431269 0.463	On request
M only	Internal	Yes	No	0.112793 0.056*	0.0833212 0.147	(Dropped)	On request
M only	Internal	No	Yes	0.109066 0.032**	0.0631235 0.195	0.0382101 0.498	On request
M only	Internal	Yes	Yes	0.0984784 0.036**	(Dropped)	(Dropped)	On request
F only	Internal	No	No	0.0738337 0.273	0.0710592 0.308	0.0283892 0.657	On request
F only	Internal	Yes	No	(Dropped)	(Dropped)	(Dropped)	On request
F only	Internal	No	Yes	0.0879482 0.174	0.0807879 0.234	0.0420571 0.503	On request
F only	Internal	Yes	Yes	0.0857581 0.142	(Dropped)	(Dropped)	On request

3.7. Analysis of mortality outside of hospital

As noted above, the HSMR parameter estimate in the PCT all cause mortality equation could be less than 0.5 (the proportion of patients dying in hospital) because HSMRs do not fully case-mix adjust for geographic variation in the demand for, and provision of, end of life care outside of hospital. This in turn alters in a non-random way the patient mix across hospitals, and undermines the assumption that all hospitals can be assumed to be treating a random selection of patients from the same pool, once the control variables in the mortality equation have been applied. The hypothesis is that in areas where out-of-hospital hospice provision is extensive, or the demand for end of life institutional care is low, HSMRs will be lower (as hospitals are dealing with fewer mortally ill patients), and more mortally ill patients will be outside of hospital, driving a negative relationship between the error terms of equations forecasting in- and out-of-hospital mortality.

To test for the existence of such a relationship, unstandardised outside-of-hospital mortality is studied for 2007-2009. Its natural logarithm is used as the dependent variable, with the natural logarithm of the HSMR rate and the previous vector of controls as the explanatory variables. Separate regressions are conducted for men and women; as before, in order to standardise for age, the full set of the relevant gender's population fractions in each 5-year age band is also included within the regression. The results are presented in Table 3.

Table 3: Parameter estimates of the influence of HSMR rates on PCT out-of-hospital mortality rates.

				HSMR coefficient p-value
Gender	**Age/sex standardisation**	**Stepwise**	**Outliers removed**	**2007-2009**
Male only	Internal	No	No	-0.420363 0.002***
Male only	Internal	Yes	No	-0.4086252 0.001***
Male only	Internal	No	Yes	-0.4302466 0.002***
Male only	Internal	Yes	Yes	-0.4553488 0.000***
Female only	Internal	No	No	-0.4748333 0.002***
Female only	Internal	Yes	No	-0.4980419 0.000***
Female only	Internal	No	Yes	-0.456779 0.002***
Female only	Internal	Yes	Yes	-0.4463307 0.001***

All coefficients are negative and statistically significant at the 1% level, with coefficient estimates around -0.4. In other words, where in-hospital deaths are 1% greater than expected, out-of-hospital deaths are 0.4% less than expected. It was shown in Table 1 that the coefficient on the HSMRs in the analysis of all-cause mortality equals around 0.1, as compared to a value of 0.5 if all extra deaths predicted by the HSMR were present in the all cause mortality rate. How far can this gap of 0.4 be explained by lower out-of-hospital deaths where HSMRs are high? Taking a typical value of -0.4 from the above regressions, and noting that 50% of deaths occur outside of hospital, around 0.2 (50% of 0.4) – or one half – of the 0.4 gap can be explained by a relationship between the HSMR and place of death.

One of the regressions fails to pass the RESET test at the 10% level and four of the regressions fail to pass the Shapiro-Wilk W test at the 10% level. Nonetheless, coefficient sizes are clearly similar and statistical significance remains very high across all eight regressions.

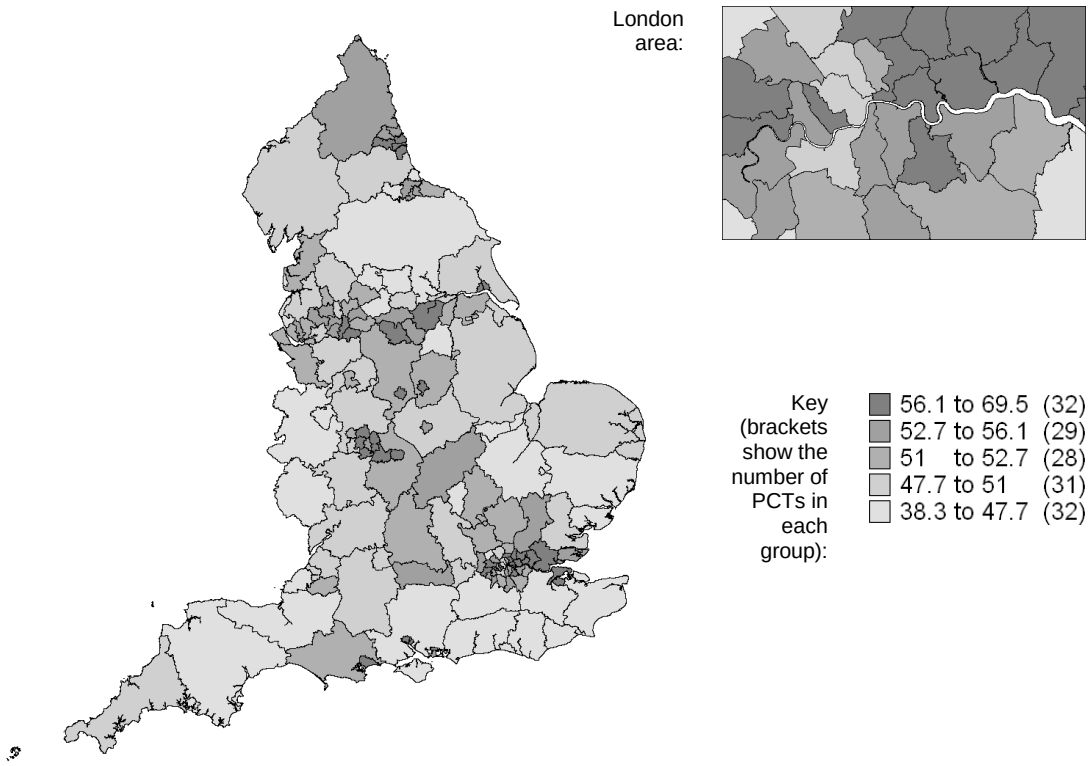
3.8. Estimating in-hospital mortality – further observations

The discussion in the previous section suggests that in those areas where in-hospital mortality rates are higher than forecast, out-of-hospital mortality rates are less than expected given the local demographic structure. This in turn suggests that the sample of patients admitted to hospitals with high HSMRs may contain disproportionately high levels of near death admissions, and which were not controlled for by the explanatory variables in the in-hospital mortality equation. It raises the difficulty of estimating in-hospital mortality models without systematically integrating a model of the likelihood that persons near death have been admitted to a hospital, or at the least including in the in-hospital mortality equation variables that are likely to influence whether near-death persons are admitted to each hospital. In England, the scale of the variation of near-death persons reaching hospital is considerable, so that modelling this source of variation in in-hospital mortality is a substantial matter. Using the data from earlier, the percentage of deaths occurring in hospital varies considerably across PCTs – from 38-67%.

This can be further explored by regressing the HSMR measure for each PCT on the demographic variables used in Table 1 and the percentage of deaths that occur in hospital. The parameter estimated on the percentage dying in hospital is 0.69 with a t-statistic of 4.23. Clearly, the HSMR statistic, based on an error term from an in-hospital mortality equation, is greater in areas where a higher proportion of deaths occur in hospital. Hence any other variable in this in-hospital mortality equation that is correlated with the likelihood of dying in hospital, holding constant demographic variables, will have a biased parameter estimate. Figure 2 provides a map of the fraction of in-hospital deaths by PCT.

There is a further statistical obstacle concerning the capability of survival amongst those reaching hospital, even within a given diagnosis category. The probability of survival for near-death patients in hospital is unlikely to be independent of the fraction of persons that die in hospital since the considerations that influence the likelihood of dying in hospital are likely to be correlated with the considerations that determine survival status when reaching hospital. Thus, for example,
(Fifth,) controlling for use of intensive hospital treatments, rural residents have greater mortality at day 1 (0.57 percentage-point), but this effect vanishes by 1 year. This effect suggests the greater availability of prehospital advanced cardiac life support systems in metropolitan areas may improve acute survival. Nonetheless, rural patients as a group have significantly worse acute and long-term outcomes as a result of their greater likelihood of receiving treatment at lower-volume, noncatheterisation hospitals, McClellan (1994)[13].

Figure 2: Percentage of deaths in hospital by PCT, 2007-2009. Darker shades denote higher rates of death in hospital.



4. Discussion

The use of regression modelling of hospital mortality, controlling for the differences in hospital patients and treatments in order to compare the safety levels of hospitals, is both widespread and potentially of very high social value. However, opinion is divided regarding how far these models measure exceptionally high/low rates of dying and might be considered a statistical artefact unconnected to all cause local area mortality. It aims firstly to show that HSMRs cannot be dismissed so lightly: higher-than-expected dying does arise in areas served by hospitals with higher HSMRs. Although the results are not statistically significant at the 5% (or even 10%) level in every case, they are consistently positive. While this strengthens the case for taking HSMRs as a serious indicator of unnecessary dying, it equally strengthens the case that unobserved local case mix variation influences both local hospital HSMRs and PCT all cause mortality.

Secondly, it shows that PCT areas that experience higher-than-expected in-hospital dying are also more likely to have lower-than-expected out-of-hospital dying. This effect accounts for part of the tendency for PCT all cause mortality to rise by less than the PCT HSMR would predict.

Thirdly, the positive correlation between the PCT HSMR and the percentage of PCT deaths in hospital suggests a failure to control adequately in the HSMR model for variation in the hospital case-mix with regard to near-death patients.

The study contributes by extending analysis of HSMRs to provide a basis in the all-cause mortality data. It relies upon observations drawn from a cross-section of 152

PCTs and, in common with other studies of HSMRs, would be strengthened by a larger cross-section. The analysis suggests that modelling the considerable variation in the percentage of area deaths that occur in hospital would improve the specification of the in-hospital mortality equation, adding explanatory power, and removing parameter bias. The findings are compatible with continued work to develop HSMRs as a safety metric.

5. References

[1] Jarman B, Gault S, Alves B, Hider A, Dolan S, Cook A, Hurwitz B and Iezzoni LI. Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999;318:1515-20

[2] Mohammed MA, Deeks J, Girling A, Rudge G, Carmalt M, Stevens A and Lilford R. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* 2009;338:b780

[3] Lilford R and Pronovost P. Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. *BMJ* 2010;340:c2016

[4] Nicholl J. Case-mix adjustment in non-randomised observational evaluations: the constant risk fallacy. *J Epidemiol Community Health* 2007;61:1010–1013

[5] Shahian DM, Wolf RE, Iezzoni LI, Kirle L, and Normand SL. Variability in the measurement of hospital-wide mortality rates. *N Engl J Med* 2010;363:2530-9

[6] Hawkes N. Patient coding and the ratings game. *BMJ* 2010;340:c2153

[7] Pitches DW, Mohammed MA and Lilford R. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Services Research* 2007;7:91

[8] Bottle A., Jarman B and Aylin P. Strengths and weaknesses of hospital standardised mortality ratios. *BMJ* 2011;342:c7116

[9] Department of Health. End of life care strategy: second annual report (2010). See http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_118810

[10] Ramsey JB. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)* 1969; 31(2):350-371.

[11] Shapiro SS. and Wilk MB. An analysis of variance test for normality. *Biometrika* 1965;52:591–611.

[12] Royston P. Approximating the Shapiro-Wilk W-Test for non-normality. *Statistics and Computing* 1992;2:117-119

[13] McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables JAMA 1994;272(11):859-866.

Annex A: Summary statistics for the final PCT-level dataset

Variable name	Mean	Std. Dev.	Min	Max
Core variables (mortality variables are per 100,000 population)				
HSMR 2005-7	100	9.45	77.6	125
HSMR 2006-8	100	9.1	76.8	130
HSMR 2007-9	99.9	8.7	75.3	125
Standardised all cause mortality 2005-7	615	76.4	381	794
Standardised all cause mortality 2006-8	602	77.6	361	785
Standardised all cause mortality 2007-9	587	76.5	364	773
Male unstandardised all cause mortality 2005-7	900	166	491	1320
Female unstandardised all cause mortality 2005-7	943	194	472	1450
Male unstandardised all cause mortality 2006-8	898	168	471	1290
Female unstandardised all cause mortality 2006-8	938	196	452	1420
Male unstandardised all cause mortality 2007-9	892	170	478	1310
Female unstandardised all cause mortality 2007-9	926	197	447	1400
Male unstandardised out of hospital mortality 2007-9	428	98.4	203	707
Female unstandardised out of hospital mortality 2007-9	450	129	166	872
Core variables (natural logarithms)				
ln(HSMR 2005-7)	4.6	0.0945	4.35	4.83
ln(HSMR 2006-8)	4.6	0.0913	4.34	4.87
ln(HSMR 2007-9)	4.6	0.0885	4.32	4.83
ln(Standardised all cause mortality 2005-7)	6.41	0.125	5.94	6.68
ln(Standardised all cause mortality 2006-8)	6.39	0.131	5.89	6.67
ln(Standardised all cause mortality 2007-9)	6.37	0.132	5.9	6.65
ln(Male Unstandardised all cause mortality 2005-7)	6.78	0.197	6.2	7.18
ln(Female unstandardised all cause mortality 2005-7)	6.82	0.235	6.11	7.26
ln(Male unstandardised all cause mortality 2006-8)	6.78	0.201	6.15	7.17
ln(Female unstandardised all cause mortality 2006-8)	6.82	0.235	6.11	7.26
ln(Male unstandardised all cause mortality 2007-9)	6.77	0.205	6.17	7.18
ln(Female unstandardised all cause mortality 2007-9)	6.81	0.239	6.1	7.24
ln(Male unstandardised out of hospital mortality 2007-9)	6.03	.243	5.31	6.56
ln(Female unstandardised out of hospital mortality 2007-9)	6.06	.318	5.11	6.77
Main control variables				
% Income deprived	17	6.54	7.06	39.2
Population density	23.9	25.7	0.8	131
% Indian	2.44	4.26	0.0767	25.7
% Pakistani and Bangladeshi	2.4	4.78	0.0475	36.1
% Black Caribbean	1.39	2.54	0.0164	12.3
% Black African	1.26	2.67	0.0307	16.1
% Chinese	0.493	0.434	0.0927	2.25
% Other Asian	0.572	0.879	0.0241	5.19
% Other Black	0.238	0.435	0.0039	2.31
% NS-SEC class 3	9.4	1.57	5.93	14.4
% NS-SEC class 4	6.66	1.71	3.7	12.1
% NS-SEC class 5	7.08	1.48	2.76	10.6
% NS-SEC class 6 and 7	20.9	4.55	9.04	30.5
% NS-SEC class 8, never worked	3.09	1.9	1.08	11.6
% NS-SEC class 8, long term unemployed	1.13	0.515	0.366	2.4
% NS-SEC unclassified, students	7.19	2.66	4.23	16.3

Variable name	Mean	Std. Dev.	Min	Max
% NS-SEC unclassified, other	17.8	3.24	11.2	25.2
Median earnings	441	62.3	343	723
Crime rate	57.6	17.6	30.9	116
% highest qualification other	6.87	1.18	3.3	9.2
% highest qualification first degree	19.9	8.48	9.7	51.5
% highest qualification 2 A-levels	8.24	1.97	5.3	15
% highest qualification 5 GCSEs	19	2.53	12	23.5
Age control variables				
% Male under 1	1.34	0.256	0.939	2.2
% Male 1 to 4	4.98	0.714	3.87	7.65
% Male 5 to 9	5.85	0.556	4.03	7.52
% Male 10 to 14	6.28	0.714	3.43	8.2
% Male 15 to 19	6.87	0.737	3.92	8.35
% Male 20 to 24	7.22	1.91	4.65	15.7
% Male 25 to 29	6.93	2.19	3.96	13.3
% Male 30 to 34	6.86	2.33	4.22	17.8
% Male 35 to 39	7.71	1.14	5.89	11.9
% Male 40 to 44	7.84	0.574	6.11	9.75
% Male 45 to 49	6.94	0.55	4.62	7.79
% Male 50 to 54	6.02	0.703	3.89	7.37
% Male 55 to 59	5.98	1.02	2.9	7.57
% Male 60 to 64	5.46	1.27	1.97	8.02
% Male 65 to 69	4.18	0.92	2.08	6.21
% Male 70 to 74	3.57	0.76	1.9	5.36
% Male 75 to 79	2.8	0.631	1.38	4.6
% Male 80 to 84	1.86	0.461	0.785	3.24
% Male over 85	1.32	0.366	0.638	2.54
% Female under 1	1.24	0.257	0.833	2.28
% Female 1 to 4	4.6	0.714	3.41	7.24
% Female 5 to 9	5.42	0.525	3.91	7.28
% Female 10 to 14	5.78	0.587	3.22	7.2
% Female 15 to 19	6.26	0.64	3.59	8.2
% Female 20 to 24	6.69	2.07	3.46	15.4
% Female 25 to 29	6.75	2.42	3.52	15.9
% Female 30 to 34	6.57	1.79	4.27	14.8
% Female 35 to 39	7.46	0.72	6.06	9.47
% Female 40 to 44	7.65	0.553	5.7	8.93
% Female 45 to 49	6.83	0.517	4.51	7.66
% Female 50 to 54	5.95	0.605	3.74	7.12
% Female 55 to 59	5.96	0.918	3.29	7.64
% Female 60 to 64	5.52	1.15	2.48	7.98
% Female 65 to 69	4.35	0.834	2.31	6.17
% Female 70 to 74	3.92	0.77	1.97	5.64
% Female 75 to 79	3.47	0.725	1.67	5.32
% Female 80 to 84	2.76	0.61	1.27	4.32
% Female over 85	2.81	0.74	1.01	5.2

Each variable has 151 observations. Values are rounded to 3 significant figures. 'Standardised' denotes direct standardisation for age & sex.

Annex B: Details of control variables used in this analysis

Variable name	Variable description	Year data relates to	Geography (number of areas covered)	Source	Web address
% income deprived	% of population classed as 'Income Deprived' in the Index of Multiple Deprivation	2005	PCT (151)	Index of Multiple Deprivation, data available from National Centre for Health Outcomes Development (NCHOD)	See (1) below
Population density	Population density, number of persons per hectare	2001	Lower-Tier Local Authority (354)	Census, data available at ONS Neighbourhood	See (2) below
% Indian	% of population in each ethnic group.	2001	PCT (303)	Census, data available from National Centre for Health Outcomes Development (NCHOD)	See (1) below
% Pakistani					
% Bangladeshi					
% Black Caribbean					
% Black African					
% Chinese					
% Other Asian					
% Other Black					
% NS-SEC class 3	% of population in each class of the National Statistics Socio-Economic Classification (NS-SEC): **Class 3** = intermediate occupations **Class 4** = small employers and own account workers **Class 5** = lower supervisory and technical, **Class 6 and 7** = semi-routine and routine **Class 8** = never worked and long term unemployed	2001	Lower-Tier Local Authority (354)	Census, data available from ONS Nomisweb	See (4) below
% NS-SEC class 4					
% NS-SEC class 5					
% NS-SEC class 6 and 7					
% NS-SEC class 8 – never worked and long term unemployed					
% NS-SEC unclassified – students and other					
Median earnings	Median gross weekly pay of full-time employees	2006/7	Lower-Tier Local Authority (354)	CLG Places (set Theme to LA and NSP National Indicators)	See (3) below
% highest qualification 5 GCSEs	% of the population whose highest qualification is: **5 GCSEs** (or equivalent) **2 A levels** (or equivalent) **A first degree** (or equivalent) **Other qualifications** (which excludes those with 'low or no qualifications', i.e. less than 5 GCSEs or equivalent)	2001	Lower-Tier Local Authority (354)	Census, data available from ONS Area Classifications	See (5) below
% highest qualification 2 A-levels					
% highest qualification first degree					
% highest qualification other					
Crime rate	Overall crime rate, recorded crime British Crime Survey Comparator	2007/8	Lower-Tier Local Authority (354)	CLG Places (set subject to Crime, then look at FTI Crime)	See (3) below

Web addresses for the above tables:

(1) http://www.nchod.nhs.uk/
(2) http://www.neighbourhood.statistics.gov.uk/dissemination/
(3) http://www.places.communities.gov.uk/
(4) https://www.nomisweb.co.uk/
(5) http://www.statistics.gov.uk/about/methodology_by_theme/area_classification/datasets.asp

Centre for
Health Service Economics
& Organisation

Skipton House
80 London Road, London SE1 6LH

Nuffield College
New Road, Oxford OX1 1NF

Tel. 020 7972 5219, www.chseo.org.uk

A Centre of the Department
of Economics, University of Oxford