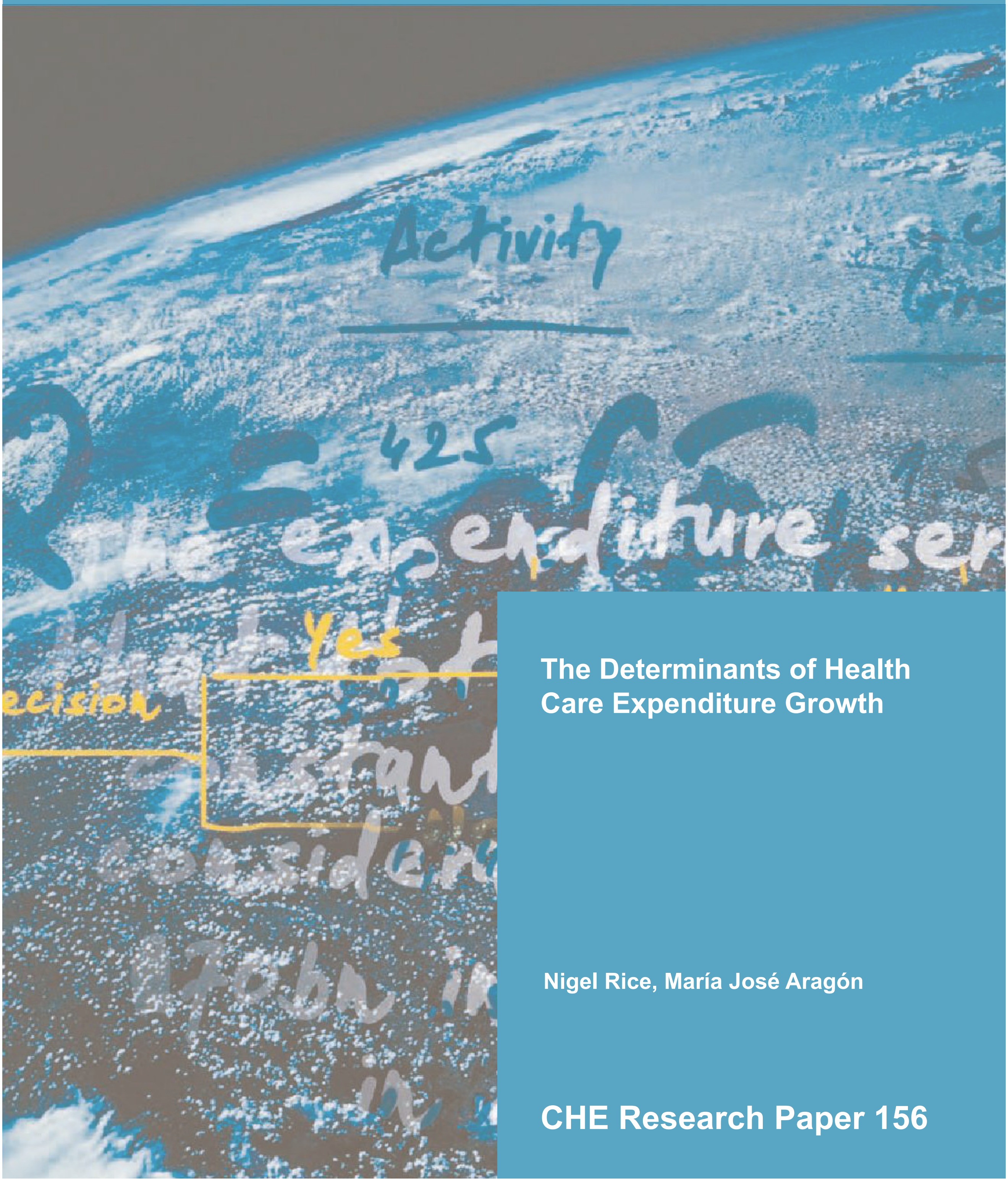
Monograph:

Rice, Nigel orcid.org/0000-0003-0312-823X and Aragon Aragon, Maria Jose Monserratt orcid.org/0000-0002-3787-6220 (2018) The determinants of health care expenditure growth. Discussion Paper. CHE Research Paper . Centre for Health Economics, University of York , York, UK.

Economics of Social and Health Care Research Unit (ESHCRU)

CHE
Centre For Health Economics

University of York



The Determinants of Health Care Expenditure Growth

Nigel Rice, María José Aragón

CHE Research Paper 156

The determinants of health care expenditure growth

Nigel Rice
María José Aragón

Centre for Health Economics, University of York, UK

July 2018

Background to series

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

This report is based on independent research commissioned and funded by the NIHR Policy Research Programme ESHCRU. The views expressed in the publication are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health, arm's length bodies or other government departments.

We are grateful for the comments provided by Hugh Gravelle and Nils Gutacker.

No ethical approval was needed.

Further copies

Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/publications/. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321405 for further details.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK
www.york.ac.uk/che

Abstract

Understanding the drivers of growth in health care expenditure is crucial for forecasting future health care requirements and to ameliorate inefficient expenditure. This paper considers the detailed breakdown of hospital inpatient expenditures across the period 2007/08 to 2014/15. Decomposition techniques are used to unpick the observed rise in expenditure into a component due to a change in the distribution of characteristics, for example, greater prevalence of morbidity, and a component due to structural changes in the impact of such characteristics on expenditures (coefficient effects, for example, due to technological change). This is undertaken at the mean using standard decomposition techniques, but also across the full distribution of expenditures to gain an understanding of where in the distribution growth and its determinants are most relevant. Decomposition at the mean indicates a larger role for a structural change in characteristics rather than a change in coefficients. A key driver is an increased prevalence of comorbidities. When considering the full distribution we observe a decrease in expenditure at the bottom of the distribution (bottom two quintiles) but increasing expenditure thereafter. The largest increases are observed at the top of the expenditure distribution. Where changes in structural characteristics dominate changes in coefficients in explaining the rise in expenditure. Increases in comorbidities (and the average number of first diagnoses) across the two periods, together with increases in non-elective long stay episodes and non-elective bed days are important drivers of expenditure increases.

JEL codes: H51; J11; I19.

Keywords: English National Health Service, Health care expenditure growth, Decomposition analysis, Drivers of expenditure.

1. Introduction and Background

The continued rise in Health Care Expenditure (HCE) relative to national income has attracted a great deal of attention and raises important questions about the sustainability of health services provision and their ability to meet population needs. Since 1978-79 while UK public spending on health rose by an average of 3.8% per year in real terms, the average growth of the economy was 2.2% a year (Licchetta and Stelmach 2016). In 2012 this equated to approximately 7.9% of GDP. The year-on-year rise in HCE is considered one of the greatest challenges to long-term fiscal sustainability (Licchetta and Stelmach 2016). Understanding the drivers of the demand for health care is critical in informing the level and distribution of future health care spending. Of particular relevance are changing demographics in an ageing population, increases in chronic conditions and comorbidities, and rising public expectations of the benefits of health care. Supply side factors including increasing relative health care costs, and the impact of technological change also heavily influence expenditure decisions. Exploring how these factors have changed over time and their relative contribution to expenditure growth is key to understanding the rise in health care expenditure and for forecasting future expenditure requirements.

This paper considers changes in hospital inpatient expenditure (a key component of HCE) over time and how these relate to changes in the determinants of such expenditure using individual-level administrative data. Using decomposition techniques in the spirit of Dormont et al. (2006) allows us to attribute changes in expenditure to a component due to structural changes, for example in demand drivers such as an ageing population, or changing morbidity characteristics and a component due to a change in the relationship between such characteristics and expenditure. The latter might arise due to changing input prices or technological progress. We further consider how these relationships and the attribution of expenditure to structural and technological change varies across the full distribution of inpatient expenditure. This allows us to look beyond mean expenditure growth to consider whether growth is driven disproportionately by expenditure at certain points across it's distribution (see de Meijer et al. (2013) for an application to the Netherlands). This is important in understanding the relative impact of structural and technological change at different parts of the distribution of expenditure. It allows the exploration of questions such as: has technological change played a relatively greater role in constraining costs at the bottom rather than the top of the expenditure distribution?

We use administrative data from Hospital Episode Statistics (HES) recorded at an individual patient level and covering all inpatient activity in England across the two financial periods 2007/08 and 2014/15. We have approximately 66,000 observations in 2007/08 and 75,000 in 2014/15, corresponding to a 1% sample of patients in each year. By matching admissions to their associated costs we are able to establish the change in hospital inpatient expenditure over the seven year period to 2014/15. HES includes demographic information, diagnoses and treatments, which allows us to explore how changes over time are determined by changes in the distribution of these characteristics (structural changes) together with changes in the expenditure response to these characteristics.

2. Heath care services in the UK

Health care delivery in the UK is predominantly provided by the publicly funded NHS.[1] Approximately 80% of health care is financed through public funds, the remainder being privately financed with this split between public and private expenditure remaining fairly constant over time. This paper focuses exclusively on NHS expenditure and provision of inpatient hospital services. The UK NHS is comprised of four separately funded systems for each of England, Northern Ireland, Scotland, and Wales with block grants used to determine the level of public funding to each of the devolved administrations. Each administration is free to decide on the level of expenditure to devote to health care and the broad functional split between primary care often termed *Family Health Services* and secondary care, often referred to as *Hospital and Community Health Services: HCHS*. The latter accounted for approximately 61% and 64% of total (NHS) health care expenditure in 2007/08 and 2014/15 respectively (Office for National Statistics 2017). HCHS covers all hospital treatments both for admitted patient care (patients who stay for at least one night as inpatients and day cases discharged on the same day as they are admitted) and outpatient care. *Family Health Services* largely covers the provision of general medical practice or ambulatory care which accounts for approximately 17% of expenditure. Prescribing by general medical practitioners and non-NHS provision represent smaller shares of around 12% and 8% respectively (Office for National Statistics 2017). Financing for the NHS is derived from conventional income and expenditure taxes, with a minority (approximately 2%) sought from patient contributions limited to a small number of services (for example, dental care, prescription charges, and eye tests).[2] With an effectively zero copayment the NHS has relied on the gate-keeping role of primary care physicians, and waiting and queuing for treatment also playing a role in regulating demand for secondary care services.

Recent reforms, mostly in England, commencing in the financial year 2003/04 have seen the adoption of fixed prices for hospital treatments, greater discretion over the use of funds by NHS hospitals and empowerment of patients through encouraging choice and 'shopping around'. These changes have been evidenced to increase hospital activity (Charlesworth et al. 2014).

2.1. Basic trends in Health Care Expenditure

Health care expenditure as a proportion of GDP has more than doubled over the past 50 years from approximately 4% in 1970 to 9% in 2012 (OECD 2017). This is illustrated in Figure 1.[3] Notable rises coincide with recessions (and a corresponding decrease in GDP) and the period in the 1990s under a government policy of year-on-year real terms increases in funding for the NHS. In the more recent past NHS expenditure as a proportion of GDP rose steeply in the two years to the end of 2009 to coincide with the great recession and then decreased during the recent period where health funding has been maintained in real terms while GDP began to rise (see Aragón et al. (2016) for details).

Focusing on the last decade, which included a large and sustained increase in health expenditures as a percentage of GDP, real NHS expenditure per capita increased over 30% for all four countries of the UK (HM Treasury 2010-2017), see Figure 2.

[1] A exception is the provision of dental care which increasingly is provided privately.
[2] Patient charges from NHS prescribed medicines in England are currently (from 01.April.2017) £8.60 per item dispensed. Charges for dental treatments vary across jurisdictions with England paying between £20.60 and £244.30 depending on the service provided. See: `https://www.gov.uk/government/speeches/nhs-prescription-charges-from-april-2017`, accessed 21.August.2017
[3] There is a break in the data in 2013, for details see OECD (2017).

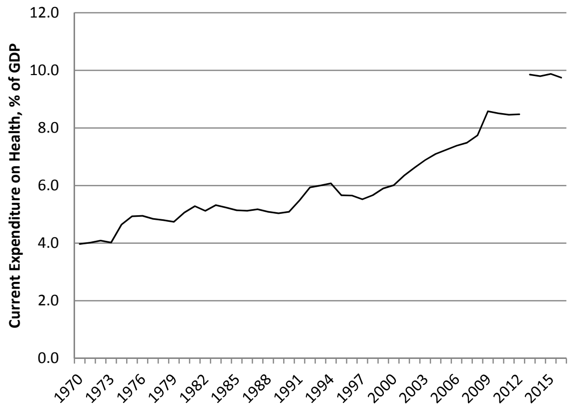


Figure 1: Total Expenditure on Health as Percentage of GDP

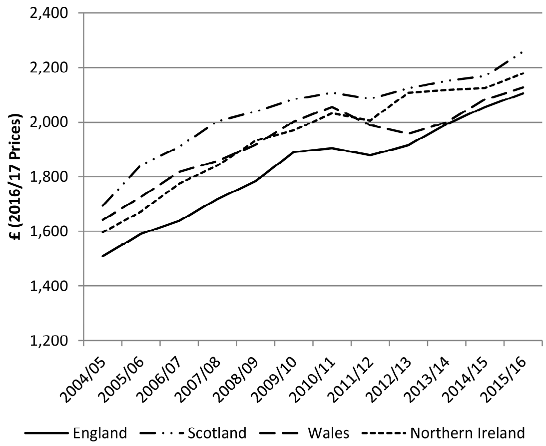


Figure 2: Health Care Expenditure per capita in the UK

Given the substantial role of hospital services in the provision of health care in England, together with the availability of inpatient hospital administrative records over a number of years which can be costed in a reasonably consistent manner, we focus our attention on this particular aspect of the NHS. This allows us to investigate changes in expenditure over the recent period between 2007/08 to 2014/15 and potential explanations for the observed changes. These are broken down between changes in expenditure due to changes in the structural distribution of patient characteristics (for example, an ageing society, changing morbidity patterns) and changes in the expenditure consequences of such characteristics (for example, due to technological change or changes in input prices).

3. Decomposition of health care expenditures

Using data for a sample of people with public health insurance, Dormont et al. (2006) use decomposition techniques to understand the sources of change in HCE between 1992 and 2000 in France. They consider expenditure on hospital care, physician consultations and prescriptions for ten-year age groups, and include both individuals with zero expenditure and patients with positive expenditure. Applying microsimulation techniques they calculate for each type of expenditure and age group, the probability of participation (i.e. having positive expenditure) and the average predicted expenditure for those with positive health care utilisation. These estimates are then used to calculate the change in expenditure between the two years, separating the effect of changes in health care provision for a given morbidity from changes in the incidence of morbidity. Their results indicate that for hospital care changes in the provision of care have contributed to increasing expenditure, while changes in morbidity have slightly reduced expenditure for people over 40 years of age. Other changes, however, also account for a significant part of the overall observed change in expenditures. Using the results from the microsimulation exercise they estimate the effects at the aggregate level, finding that for hospital care changes in morbidity have reduced expenditure, however they are partially offset by the effect of the changes in the provision of care. Changes in patients' age explain only a small proportion of the change in hospital expenditure.

More recently, de Meijer et al. (2013) analyse the change in the distribution of HCE between 1998 and 2004 in the Netherlands (during this period, in 2001, hospitals' budgets where relaxed to address increasing waiting lists) using a combination of insurance claims, hospital discharges and mortality data. They divide overall HCE into two components: hospital care and pharmaceuticals, and analyse whether changes in hospital activity are reflected in pharmaceutical expenditure. They decompose the change in the distribution of HCE using the methodology proposed by Chernozhukov et al. (2012), which decomposes the change into changes in its structural determinants (e.g. population age) and changes in the impact the determinants have on HCE (for example, brought about via technology improvements). Importantly the method allows this decomposition to take place across the full distribution of HCEs such that a comparison of the determinants can be made at different points in the distribution. More standard approaches to decomposition analysis (e.g. Oaxaca (1973) and Blinder (1973)) focus exclusively on decomposing changes in outcomes solely at the mean. The results of de Meijer et al. (2013) indicate that changes in HCE determinants and changes in their impact differ at the bottom of the expenditure distribution compared to the top. For individuals with positive expenditure, both determinants (characteristics) and their impact contribute to increase expenditure throughout the distribution. However, their relative importance varies. At the bottom of the expenditure distribution both factors contribute equally, while at the top, the contribution of the change in the impact is around one and a half times the contribution of the change in determinants. For hospital care, the growth in expenditure over the period is mostly explained by changes in structural determinants.

As part of calculating HCE projections, the OECD estimated expenditure growth for member states over the period 1970-2002 (OECD 2006). For the UK, they found that public health expenditure grew on average 3.8% per year, of which it was estimated that 0.1% was due to demographic effects and 2.1% to income effects (assuming health care is a normal good and an income elasticity of health expenditure equal to 1). The remaining 1.5% was ascribed to a residual effect, attributed to changes in medical technology and changes in relative prices. As improvements in medical technology can reduce cost and accordingly the relative price of health services, then depending on the elasticity of demand for health care, this may lead to increases or decreases in the demand for these services and associated expenditure.

4. Methods

4.1. Oaxaca-Blinder Decomposition

Oaxaca (1973) and Blinder (1973) analysed the wages of different groups (male/women, white/black) in the US, to investigate the role of discrimination in the observed differentials. Both authors, separately, developed methods to decompose the average wage differential into a part explained by observed worker characteristics (job tenure, ability, etc.) and a part due to discrimination. The latter can be thought of as the differential impact of a given characteristic, say ability, on wages observed between men and women. In the absence of discrimination, both men and women would be expected to be paid equally for a given level of ability, with all else held equal. Accordingly, the approach aims at decomposing wages into a discriminatory effect, arising from differential pay between two groups with the same characteristics, and a composition effect, where differences in wages arise from legitimate differences in the groups. Applying the Oaxaca-Blinder methodology, we decompose the difference in average hospital expenditures observed between two financial years, and consider its determinants due to structural changes in characteristics of patients and care provision, and changes in the expenditure response to those characteristics. The former can be thought of as picking up changes in demographic and morbidity characteristics, and the latter due to technological changes or changes in input prices. We assume individual hospital expenditures, Y_{it} in a given financial year can be estimated as:

$$Y_{it} = X_{it}'\beta_t + \epsilon_{it} \tag{1}$$

where X_{it} includes patient (age, sex) and hospital admission (number, length, diagnoses recorded) characteristics and a constant. β is a vector of parameters and ϵ_{it} is an idiosyncratic error. In our model, t denotes the time period under investigation - either 2007/08 or 2014/15. Note that these two years represent repeated cross-sections of health care inpatient users rather than a panel of a fixed set of individuals observed over time.

The difference between the average individual hospital costs in two different financial years can be calculated using Equation (1) evaluated on the mean for each period, assuming ϵ_{it} has mean zero $\forall t$. Assuming two periods denoted $t = 0$ and $t = 1$, and supressing the individual subscript for ease of notation, the equation can be rearranged to show the components of the difference as follows:[4]

$$\begin{aligned} E(Y_1) - E(Y_0) &= E(X_1'\beta_1 + \epsilon_1) - E(X_0'\beta_0 + \epsilon_0) \\ &= E(X_1)'\beta_1 - E(X_0)'\beta_0 \\ &= \underbrace{\{E(X_1) - E(X_0)\}'\beta_0}_{A} + \underbrace{E(X_0)'(\beta_1 - \beta_0)}_{B} + \underbrace{\{E(X_1) - E(X_0)\}'(\beta_1 - \beta_0)}_{C} \end{aligned} \tag{2}$$

Equation (2) illustrates the decomposition of the difference in the mean outcome between the two periods $(t = 0$ and $t = 1)$. The component labelled A shows the contribution due to a change in characteristics of patients $(X_0$ and $X_1)$ across the two periods assuming the relationship between characteristics and expenditure are observed at time $t = 0$ (herein referred to as the "characteristics effect"). The component

[4] We follow the rearrangement proposed by Jann (2008).

labelled B is the contribution to the change in outcomes due to a change in relationship between characteristic X_0 and the outcome across the two periods, captured by the difference in coefficients $(\beta_1 - \beta_0)$. That is, $(\beta_1 - \beta_0)$ indicates how, for example, a change in the efficiency of treating patients with characteristics X_0 informs the change in mean expenditures across the two time periods. The final component, C, is the effect of an interaction term between A and B. The focus of our interest lies in the relative contributions of components A and B in explaining the change in expenditures across the two periods.

The decomposition is straightforward for cardinal variables. However, for categorical variables that require a set of dummy variables to be defined together with a reference category, Oaxaca and Ransom (1999) show that the standard decomposition is dependent on the chosen reference category. Accordingly, the decomposition into effects of characteristics and coefficients is not invariant to the reference category. This is clearly problematic for empirical analyses where dummy variables are often specified. The problem arises due to a lack of agreement on which category should form the reference group. Yun (2005) proposes a simple solution to this problem based on the idea that the characteristics and coefficients effects for each outcome of the categorical variable can be computed as the average effect for that outcome when the effects are calculated by alternating the reference group. That is, multiple regressions and decompositions can be estimated, each specifying a different reference group and by taking averages of appropriate terms across the regressions the characteristic and coefficient effects for each category can be identified. Since changing the reference category for dummy variables impacts on the constant term, the above approach will also lead to a modified estimate of the constant. In practice the characteristics and coefficient effects can be derived from a single regression without the need to estimate multiple specifications each with a different reference category. If we think of X_{it} in equation (1) as a discrete variable with three categories $(k = 1, \ldots, 3)$, then the model can be extended to:

$$Y_{it} = \alpha_t + \bar{\beta}_t + \sum_{k=1}^{3} X_{kit} \left(\beta_{kt} - \bar{\beta}_{kt} \right) + \epsilon_{it} \tag{3}$$

where $\bar{\beta}_{kt} = (\beta_{1t} + \beta_{2t} + \beta_{3t})/3$ and $\beta_{1t} = 0$, and we explicitly note the constant α_t. Equation (3) is the normalised regression where the estimate is simply the average of three sets of estimates where the reference group is varied (see Yun (2005) for further discussion). The overall constant term is given by $\alpha_t + \bar{\beta}_{it}$. Decomposition analyses follows directly on the specification in model (3).

The approach generalises to the situation with multiple categorical variables. Due to the approach averaging across specifications where the reference category is changed, this leads to effects that are equal in magnitude, but with opposing directions for a single dummy variable containing two categories.

We perform the analysis using the Oaxaca-Blinder decomposition developed for Stata by Jann (2008) in Stata 14 (StataCorp. 2015). For categorical variables, for example sex, we use the *normalize()* option which will report the decomposition results for all categories.

In section 6.1 we decompose the natural logarithm of expenditures across the two years of data, such that (2) can be represented as $E(lnY_1) - E(lnY_2)$. Accordingly decomposition is undertaken on a logarithmic scale, but results are retransformed (exponentiated) to the original scale prior to reporting. Retransforming the data allows for a more intuitive interpretation of the decomposition effects. Since this involves a ratio of expenditures across the period it informs of the percentage increase (or decrease) in expenditure due to a change in coefficients or change in structural characteristics.

4.2. Counterfactual Decomposition

The Oaxaca-Blinder decomposition described in Section 4.1 focuses on decomposing the average differences between groups. The key decompositional effects are differences between observed and counterfactual distributions evaluated at the mean. The counterfactual can be thought of as the result either of a change in the distribution of a set of characteristics predictive of the outcome, or a change in the relationship between the characteristics and the outcome. Essentially the counterfactual represents the 'what if' scenarios - for example, component A in Equation (2) reflects the impact of a change in the distribution of characteristics, X, evaluated using the estimated relationship with the outcome (β_0) observed in the first of the two time periods, that is at $t = 0$. The counterfatual to the observed relationship at time $t = 0$ (which estimates the relationship between X_0 and Y_0 as β_0) is the impact of characteristics observed in period $t = 1$, that is, X_1 applied to the relationship with Y_0 observed in period $t = 0$, that is, β_0. Accordingly, $X_1\beta_0$ represents the counterfactual 'what if' scenario at $t = 1$ for the estimated relationship $X_0\beta_0$ observed at time $t = 0$. That is, by applying the relationship between outcomes and characteristics at time $t = 0$ what would the outcome be if the set of characteristics were those observed at time $t = 1$. Whereas the Oaxaca-Blinder approach evaluates the counterfactual at the mean of the outcome, other techniques consider the difference between groups across the full distribution of outcomes. Fortin et al. (2011) provide an overview of methods to decompose the difference between two groups (or two time periods) into different explanatory factors.

To investigate the change in HCE across its full distribution (rather than solely at the mean) and to decompose this into changes in characteristics and changes in the impact of characteristics we follow the distributional regression approach developed by Chernozhukov et al. (2013). They propose a methodology to undertake counterfactual analysis, that is, to evaluate the effect of changes in the (marginal) distribution of an outcome variable Y given a set of covariates X, either via a change in the distribution of X or a change in the relationship between X and Y (the conditional distribution of Y given X $(F_{Y|X})$). In general for a given outcome, Y, and set of characteristics, X, we are interested in the effect on the marginal distribution of the outcome (F_Y) for a change in: (i) the marginal distribution of the characteristics (F_X) holding fixed the conditional distribution of the outcome $(F_{Y|X})$; (ii) the conditional distribution of the outcome $(F_{Y|X})$ holding the marginal distribution of characteristics (F_X) fixed. To address these effects requires the estimation of counterfactual distributions.

Chernozhukov et al. (2013) propose a decomposition method based on distributional regression (for example, see Foresi and Peracchi (1995)). More formally, we can define the marginal distribution of characteristics in time period j $(j = 0, 1)$ as $F_{X_{(j)}}(x)$, and the conditional distribution of the outcome given characteristics X as $F_{Y_{(j)}}(y|x)$. The insight behind the decomposition comes from the fact that the marginal distribution of the outcome, $F_{Y_{(j)}}$, is equivalent to it's conditional distribution integrated over the distribution of covariates. We can represent this as: $F_{Y_{(j)}}(y) = \int F_{Y_{(j)}}(y|x)dF_{X_{(j)}}(x)$. Counterfactual analysis follows directly by generalising the above, such that, for example, $F_{Y_{(j|k)}}(y)$ represents the marginal distribution of the outcome in preriod j based on the distribution of characteristics measured in period k $(k = 0, 1)$. Accordingly, we have: $F_{Y_{(j|k)}}(y) = \int F_{Y_{(j)}}(y|x)dF_{X_{(k)}}(x)$. For example, the hypothetical distribution of HCE in period $t = 0$ if the relationship between characteristics and outcomes remained as observed in time period $t = 0$, but applied to the distribution of characteristics observed in period $t = 1$ is given by: $F_{Y_{(0|1)}}(y) = \int F_{Y_{(0)}}(y|x)dF_{X_{(1)}}(x)$. The approach allows comparison between marginal and counterfactual distributions of outcomes and their decomposition. To simplify notation, we drop the parentheses such that $F_{Y_{(0|0)}}$ is the observed distribution of expenditures in period $t = 0$ and $F_{Y_{(0|1)}}$ is the counterfactual distribution of expenditures assuming patients had characteristics observed in period $t = 1$; and similarly for $F_{Y_{(1|1)}}$ and $F_{Y_{(1|0)}}$. In each (financial) year, we observe hospital costs Y_t

and patient and hospital characteristics X_t. The observed distributions of hospital costs are $F_{Y_{(0|0)}}$ and $F_{Y_{(1|1)}}$, and their difference can be decomposed into differences due to differences in the coefficients (1) and to differences in characteristics (2) as:

$$F_{Y_{(1|1)}} - F_{Y_{(0|0)}} = \underbrace{\left[F_{Y_{(1|1)}} - F_{Y_{(0|1)}}\right]}_{(1)} + \underbrace{\left[F_{Y_{(0|1)}} - F_{Y_{(0|0)}}\right]}_{(2)} \quad (4)$$

Implementation of the approach requires estimation of the conditional distribution of outcomes and of the marginal distribution of characteristics. It is assumed that the conditional distributions, $F_{Y_{(j|k)}}$, are well defined where there is common support across the distribution of characteristics, X_0 and X_1. In our application this assumes that the support of patient characteristics in the initial year includes that of the final year, $\chi_1 \subseteq \chi_0$. This appears a reasonable assumption.

We perform the analysis using the decomposition commands developed for Stata by Chernozhukov et al. (2013) in Stata 14 (StataCorp. 2015).

5. Data

We focus on the population of inpatient hospital users, that is, people who attended hospital at least once, in either of the two financial years 2007/08 and 2014/15. There are two main sources of data: the Admitted Patient Care part of the Hospital Episode Statistics (APC-HES) (NHS Digital 2015; The Health and Social Care Information Centre 2009) and the NHS Reference Costs (RC) (Department of Health 2015; Department of Health 2009). APC-HES records hospital activity as *episodes*, periods under the care of one consultant. The information recorded for each episode includes: start and end dates, age and sex of the patient, type of admission and diagnoses. We can identify the Healthcare Resource Group(s) (HRG, NHS equivalent to DRGs) associated with each episode using the episode information from APC-HES and the HRG RC Grouper for the corresponding financial year (National Casemix Office 2015; The Casemix Service 2008), which also provides the number of excess bed days (days above a trimpoint defined to improve comparability of episodes by separating unusually long episodes). All episodes have a core HRG, which corresponds to the primary reason for admission or treatment, and they may also have unbundled HRG(s), which account for activity with significant costs (e.g. Chemotherapy, Radiotherapy, Specialist Palliative Care, High Cost Drugs). This separation (previously each episode had only one HRG) was introduced so HRGs represent activity and costs more accurately (National Casemix Office 2015).

The cost associated with each HRG is available in the RC, which report a (weighted) national average of the cost of each HRG based on the costs reported by NHS providers (Department of Health 2014). Each HRG can have more than one cost, depending on the settings (inpatient, day cases, non-elective) in which it has activity.

Using RCs rather than the National Tariff (NT, see Monitor and NHS England (2013)) allows us to compare years further apart. The version of the HRG classification used for RC and NT has changed over time, with RC adopting a new version a few years earlier; RC started using HRG4 in 2006/07[5] (Department of Health 2008) and NT in 2009/10 (DH PbR Team 2009). This greater span is required to ensure sufficient time has elapsed between periods under comparison to allow for meaningful change to occur.

To calculate the cost of each episode we use both the core and the unbundled HRGs. The core HRG is matched with its cost in the relevant setting, i.e. we use the elective cost for elective inpatient admissions, the day case cost for day cases, and the non-elective cost (short- or long stay, depending on the length of the episode) for the rest of the episodes; the cost for excess bed days, if these are present, is also based on the core HRG. And the unbundled HRG(s) (if any) are matched with their overall national average (without distinguishing between elective and non-elective).

Our dependant variable is the total cost per patient in 2007/08 and 2014/15, so we calculate its total per patient in each financial year, using the RC for that year. We consider only episodes with a positive cost. Accordingly, our data set considers a cross-section of inpatient hospital users in 2007/08 and a further cross-section of hospital users in 2014/15. Some individuals may be present in both data sets should they have had an inpatient episode in both financial years.

The variation in costs we observe is a combination of changes in the costs of hospital activity and changes in that activity. To take into account the effect of inflation and make the costs comparable, we

[5] We did not use 2006/07 as it was not possible to use the RC Grouper for that year.

use the Department of Health's Pay & Price Series (Department of Health 2016) to deflate the costs in 2014/15 (i.e. express them in £ of 2007/08). Therefore, the change we observe is not driven by costs' inflation, but likely to reflect changes in hospital activity, such as technological change and complexity of cases.

As explanatory variables we consider patient characteristics, clinical information and activity levels in each financial year. In terms of patients characteristics, we consider the age at the beginning of the financial year and the sex of the patient. Using the (ICD-10) diagnoses recorded in each episode we create two sets of indicators based on the diagnosis groups defined in the International Shortlist for Hospital Morbidity Tabulation (ISHMT) (World Health Organization n.d.).[6] The first set of indicators considers only the first diagnosis; the second set considers the following four diagnoses. Accordingly, the first set of indicators measure the different reasons a patient is admitted to hospital (multiple admissions with the same primary diagnosis will have only one indicator equal to one) while the second set of indicators measures comorbidities the patient had during the financial year,[7] we use this second set of indicators to calculate the total number of comorbidities. In terms of activity we calculate totals for the number of episodes (total, elective/non-elective, short(day cases)/long) and number of days admitted to hospital (total and elective/non-elective). To account for provider effects we need to associate each patient to one provider, since it is possible for patients to be admitted in different providers throughout the year, we use the one where the patient had most episodes.

There are approximately 6.6 million hospital patients in the financial year 2007/08 and 7.5 million in 2014/15. To speed up computational time we take a random 1% sample from each year. Accordingly, the sample for analysis consist of 66,079 in 2007/08 and 75,343 in 2014/15.

5.1. Descriptive Statistics of the Sample

Table 1 shows the descriptive statistics of our sample: 1% of the patients in each financial year, after excluding those who have more than 365 days admitted to hospital in the year.[8] The average cost per patient increased from £2,965.27 in 2007/08 to £3,606.87 in 2014/15; an increase of 21.6%. Patients were, on average, slightly older while the proportion of males remained constant (and less than the proportion of females) across the two periods. There were more diagnoses recorded in 2014/15 both in the average number of first diagnoses and the number of comorbidities recorded.[9] Patients had more episodes in total in 2014/15 compared to 2007/08. While the number of elective episodes remains constant over the period, its composition changes - daycases increase and inpatients decrease. For both short and long stay non-electives, the average number of episodes increased. Elective hospital stays have become shorter while non-elective stays were longer in 2014/15 compared to 2007/08. The

[6] ISHMT has 130 diagnosis groups listed under 20 chapters; the results on the main part use the 20 chapters (see Appendix A for their description) and results using all 130 diagnosis group can be found in Appendix B

[7] We focus on the first five diagnoses to avoid any potential issues arising from changes in the recording of diagnoses between the two years we consider. The recording of diagnoses in HES changed between the two financial years we consider, in 2007/08 is was possible to record up to 14 diagnoses and by 2014/15 this number had increased to 20. Additionally, in 2011 the HSCIC (now NHS Digital) issued a list of diagnoses that are always considered to be clinically relevant and therefore should always be recorded (Health and Social Care Information Centre 2017).

[8] HES records finished episodes, therefore they are recorded on the financial year they end. Patients with long hospital stays can appear to have more than 365 days in hospital in a financial year, there were 2,949 such observations in the data before extracting the 1% sample.

[9] The average number of first diagnoses is above one due to some patients having multiple admissions during the year with different first diagnosis.

proportion of patients in the different types of providers remains fairly constant across the period except for a shift towards teaching providers[10] and Other Providers.[11]

The total number of patients increases between 2007/08 and 2014/15 by around 12% (in our sample the increase is of 14%), while the population in England increased by less than 6% over the same period (Office for National Statistics 2015; Office for National Statistics 2013). Accordingly, the proportion of people admitted to hospital increased over the period. In terms of capacity, the total number of available beds (overnight and day-only) increased by 2% (NHS England n.d.[b]; NHS England n.d.[a]).

[10] In 2008/09 (the closest year to 2007/08 for which we have data) there were 25 teaching hospitals in England compared to 29 in 2013/14.

[11] Other Providers consist largely of non-acute trusts. In 2007/08 these included mental health (MH) trusts and primary care organisations (PCOs). In 2014/15, these included MH Trusts and community trusts together with a merger in October 2014 between Ealing Hospital NHS Trust and The North West London Hospitals NHS Trust, which created one of the largest integrated acute and community care trusts in the country.

Table 1: Descriptive Statistics

	2007/08		2014/15	
	Mean	StdDev	Mean	StdDev
Cost (in £ of 2007/08)	2,965.27	5,277.25	3,606.87	6,506.78
Demographics				
Age	49.97	25.35	51.64	25.41
Male	0.4661	0.4989	0.4655	0.4988
Morbidity				
Number of different 1st diagnoses				
based in 20 ISHMT chapters	1.26	0.61	1.33	0.70
based in 130 ISHMT groups	1.33	0.74	1.41	0.85
Number of Comorbidities (based on 130 ISHMT groups)	1.20	1.60	1.88	1.96
Activity				
Total Episodes	2.03	4.49	2.17	2.84
Elective Episodes	1.09	4.17	1.09	2.26
Daycases	0.68	1.64	0.81	1.82
Inpatient	0.41	3.76	0.29	1.24
Non-Elective Episodes	0.93	1.56	1.08	1.83
Short Stay	0.18	0.49	0.22	0.53
Long Stay	0.75	1.46	0.86	1.73
Bed Days	5.01	13.91	5.10	15.72
Elective	0.95	4.69	0.84	6.71
Non-Elective	4.07	12.82	4.27	13.93
Provider Type				
Large Provider	0.3612	0.4803	0.3153	0.4646
Medium Provider	0.2780	0.4480	0.2573	0.4371
Small Provider	0.1100	0.3129	0.0996	0.2995
Specialist Provider	0.0281	0.1651	0.0278	0.1645
Teaching Provider	0.2207	0.4147	0.2802	0.4491
Other Provider	0.0020	0.0441	0.0198	0.1393
Number of Observations	66,079		75,343	

6. Results

6.1. Decomposition at the mean of expenditures

Table 2 shows the results of the Oaxaca decomposition of the change in expenditure between 2007/08 and 2014/15. These are provided for model specifications that differ in the number of characteristics used as control variables. The list of characteristics used for each model is shown in the bottom panel of the table. The dependant variable in all regressions is the logarithm of expenditure per patient in the relevant financial year. However, the results have been transformed to the original scale, so the (geometric) mean[12] of the cost in each year is expressed in pounds (deflated to 2007/08) and the estimated difference (and its decomposition) is expressed as a ratio. For example, the geometric mean of expenditure in 2007/08 is £1,561.10 and £1,673.97 in 2014/15.[13] This represents a growth in expenditure (on the log scale) of 7.23% across the period. This difference is decomposed into an effect due to a change in characteristics (labelled Characteristics) and an effect due to a change in the relationship between characteristics and expenditure (labelled Coefficients). In model specification (1) the increase in costs due to changes in patients and hospital admission characteristics is 2.1$ (characteristics ratio = 1.0209). The change due to how characteristics relate to expenditure is 5.2% (coefficient ratio = 1.0524). The interaction between characteristics and coefficients (how the characteristics relate to expenditure) explains a further 0.2% of expenditure.

All estimates reported in the top panel of Table 2 are significant at 1%. Exceptions to this are specification (7) for the effect of a change in coefficients and specifications (4) and (5) for the interaction term, which have p-values > 0.10. In the first two specifications, the overall change is driven by the change in coefficients. Once first diagnosis is included (specification (3) onwards) changes in characteristics become the main component of the overall change in expenditure. The inclusion of comorbidities (specification (4) onwards) changes not only the relative sizes of the effect of characteristics and coefficients, but also the sign of the effect of coefficients.

We focus in detail on the results for the most basic and most complex specifications (1) and (7). Model (1) considers only patient demographic characteristics (age and sex) as explanatory variables. Table 3 shows the full results for this specification. The first two columns present regression results for each of the two financial years. Decomposition results are reported in the final three columns. The first row reproduces the decomposition results reported in Table 2 (including significance levels and standard errors). The second row shows the same decomposition but excluding the constant term, which is only present in the results for the Coefficients. The regression results for the two financial years (first two columns in Table 3) are broadly similar. In both years women, on average, attract lower cost and costs generally increase with age, particularly from age 30-40 upwards.

The decomposition results explain the change in expenditures evaluated at the mean. Of the observed 7.2% increase in expenditure across the two periods, 2% can be explained by a change in characteristics (demographic variables) and 5.2% by a change in the relationship between demographic variables and expenditure. Note that this last effect includes a constant term; if this term is excluded, the effect of the changes in characteristics becomes smaller and changes sign (-1%). This specification, given it's

[12] This is not the same as reported in the descriptives statistics in Table 1.

[13] The geometric mean indicates the central tendency by calculating the n^{th} root of the product of n numbers. This differs from the arithmetic mean which calculates the average of a sum of n numbers. Another way of expressing the geometric mean is the exponential of the arithmetic mean of the logarithm of the n numbers. Unless all numbers are equivalent, in which case the two means coincide, the geometric mean is always less than the arithmetic mean.

Table 2: Blinder-Oaxaca Decomposition of Hospital Costs Difference between 2007/08 and 2014/15

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cost in 2014/15	1,673.97	1,673.97	1,673.97	1,673.97	1,673.97	1,673.97	1,673.97
Cost in 2007/08	1,561.10	1,561.10	1,561.10	1,561.10	1,561.10	1,561.10	1,561.10
Difference	1.0723	1.0723	1.0723	1.0723	1.0723	1.0723	1.0723
Characteristics	1.0209	1.0232	1.0426	1.1149	1.1141	1.1027	1.0889
Coefficients	1.0524	1.0336	1.0228	0.9647	0.9602	0.9632	0.9933
Interaction	0.9981	1.0139	1.0056	0.9970	1.0024	1.0096	0.9913
Variables:							
Demographics							
Age	X	X	X	X	X	X	X
Sex	X	X	X	X	X	X	X
Activity							
Total Episodes		X	X	X	X		
Elective Episodes						X	
Daycases							X
Inpatient							X
Non-Elective Episodes						X	
Short Stay							X
Long Stay							X
Bed Days		X	X	X	X		
Elective						X	X
Non-Elective						X	X
Morbidity							
First Diagnosis			X	X	X	X	X
Number of Comorbidities				X	X	X	X
Provider							
Provider Type					X	X	X
Number of Observations 2007/08				66,079			
Number of Observations 2014/15				75,343			

simplicity, is likely to mask other relevant characteristics related both to changes in expenditure and to age.

A detailed breakdown of the decomposition results by individual covariates and groups of covariates is also provided. For categorical variables, these are estimated using the *normalize()* option and report results for all categories of a particular categorical variable, together with the overall effect of the variable (between dashed lines). For example, in the case of sex, there are separate coefficients for males and females. As the results are shown as a ratio, a ratio equal to one indicates no relationship with the change in costs for that particular change in characteristic (or change in coefficient). The results show that sex is not significantly related to the observed change in cost. Neither the distribution of males and females, nor the relationship between sex and expenditure contribute to explaining the observed change in expenditures across the period. That is, the ratio both for the effect of a change in characteristic and a change in coefficient are close to unity.

Age, on the other hand, is significant for the change attributed to characteristics, coefficients and the interaction term. The change in expenditure attributed to a change in characteristics is positively correlated (ratios greater than one) for age groups from [10, 20) to [80, 120). This indicates that the rise

Table 3: Blinder-Oaxaca Decomposition of Hospital Costs Difference between 2007/08 and 2014/15 - Specification (1)

			Decomposition (as ratio)		
	2007/08	2014/15	Characteristics	Coefficients	Interaction
Decomposition (as ratio)			1.0209***	1.0524***	0.9981***
			(0.0017)	(0.0060)	(0.0007)
Decomposition (as ratio) excluding Constant			1.0209	0.9882	0.9981
Male	Reference Category		1.0000	0.9971	1.0000
			(0.0001)	(0.0026)	(0.0000)
Female	-0.0675***	-0.0553***	1.0000	1.0033	1.0000
	(0.0079)	(0.0083)	(0.0001)	(0.0031)	(0.0000)
Sex - total effect			1.0000	1.0004	1.0000
			(0.0002)	(0.0004)	(0.0000)
Age Groups					
[0, 5)	Reference Category		0.9999	1.0008	1.0000
			(0.0003)	(0.0014)	(0.0000)
[5, 10)	-0.0317	-0.0262	1.0000	1.0005	1.0000
	(0.0290)	(0.0304)	(0.0002)	(0.0008)	(0.0000)
[10, 20)	0.0641***	0.1482***	1.0020***	1.0061***	0.9990***
	(0.0219)	(0.0242)	(0.0003)	(0.0014)	(0.0003)
[20, 30)	-0.0207	-0.0096	1.0017***	1.0021	0.9999
	(0.0201)	(0.0215)	(0.0004)	(0.0017)	(0.0001)
[30, 40)	0.0481**	0.0401*	1.0035***	1.0005	0.9999
	(0.0196)	(0.0214)	(0.0004)	(0.0018)	(0.0003)
[40, 50)	0.1701***	0.1301***	1.0004**	0.9967*	1.0001
	(0.0191)	(0.0202)	(0.0001)	(0.0019)	(0.0001)
[50, 60)	0.3244***	0.1816***	1.0006***	0.9839***	0.9990***
	(0.0190)	(0.0197)	(0.0002)	(0.0019)	(0.0003)
[60, 70)	0.4704***	0.3765***	1.0027***	0.9884***	0.9990***
	(0.0185)	(0.0192)	(0.0004)	(0.0021)	(0.0002)
[70, 80)	0.6129***	0.5816***	1.0020***	0.9973	0.9999
	(0.0185)	(0.0194)	(0.0007)	(0.0021)	(0.0001)
[80, 120)	0.8789***	0.9694***	1.0080***	1.0125***	1.0013***
	(0.0191)	(0.0197)	(0.0011)	(0.0019)	(0.0003)
Age - total effect			1.0209***	0.9885***	0.9981***
			(0.0017)	(0.0031)	(0.0007)
Constant	7.0603***	7.1288***	-	1.0642***	-
	(0.0157)	(0.0165)		(0.0069)	
Adjusted R-squared	0.0811	0.0740			
N	66,079	75,343		141,422	

Standard errors in parenthesis. ***, **, * represent 1%, 5% and 10% significance, respectively.

in expenditure from 2007/08 to 2014/15 can, in part, be explained by a change in the distribution in the age of patients, particularly in the older age groups (60 upwards).[14] While the ratio for the relationship between age and expenditure is positive (coefficient ratio greater than 1) for age groups [10, 20) and [80, 120) it is negative (ratio less than 1) for age groups [40, 70). Taken together (results between dashed lines), the change in the distribution of age contributes 2.1% to the overall change in expenditures, whilst the change in the associated coefficients contributes -1.2% to the overall change in expenditure. Accordingly, in general, the change in the distribution of age has contributed to increasing expenditure across the two periods (an ageing population), while the change in coefficients has contributed to a decrease in expenditure (patients more efficiently treated). For the very old (80 years plus), both a change in the distribution (more elderly patients) and a change in the coefficients both contribute to the observed growth in expenditures.

The largest single contribution to the change in expenditure is due to the change in coefficients attached to the constant term (rather than individual demographic variables). In itself, this contributes 6.4% (ratio = 1.0642) to the overall effects of the impact of coefficients (5.2%) and is offset by an overall decrease in the change in coefficients on the demographic variables (-1.2% for age, sex is not significant). The constant term estimates the mean of expenditure for the reference category in the regression model (male, aged [0,5)) around which deviations are estimated for other sex and age categories. The contribution of the change in coefficients for the constant reflects an increase in average expenditure from 2007/08 to 2014/15. However, while the constant reflects the average expenditure in the reference category, much of this remains unexplained. Clearly, for model (1) the change in the constant over the period is substantial reflecting characteristics of patients or the provision of inpatient care not reflected in the set of patient demographic variables.

Model (7) is the most detailed specification and contains patient characteristics (age and sex), type of episode (elective - day case, inpatient case; non-elective - short stay, long stay) and associated bed days, diagnosis, the number of comorbidities and provider type as explanatory variables. Table 4 presents detailed results for this model using 20 chapters of the ISHMT classification for diagnoses (results using 130 diagnosis groups can be found in Appendix B) and Figure 3 summarizes these results.

The regression results presented in the first two columns in Table 4 show that conditional on information on episodes, diagnoses and provider type, the impact of sex and age is diminished compared to the more simple specification presented in Table 3. Indeed, the impact of sex on expenditures becomes non-significant. The coefficients on age, particularly for age groups of 40 years and above are much smaller in absolute terms than the corresponding coefficients presented in Table 3. Conditional on the other explanatory variables, in 2014/15 we observe negative and statistically significant coefficients for some age groups. This is not unexpected. Howdon and Rice (2017) show that time-to-death is a stronger predictor of health care expenditures than age, but that time-to-death is itself a proxy for morbidity. Conditional on measures of morbidity (via diagnoses) it is, therefore, not surprising that age plays a less crucial role in predicting expenditures. In model (1) the effects ascribed to age are reflecting, in part, unspecified correlates of age which will include morbidity.

[14] Recall from Table 1 that the mean age of patients increases from approximately 50 years in 2004/05 to approximately 52 years in 2014/15.

Table 4: Blinder-Oaxaca Decomposition of Hospital Costs Difference between 2007/08 and 2014/15 - Specification (7)

			Decomposition (as ratio)		
	2007/08	2014/15	Characteristics	Coefficients	Interaction
Decomposition (as ratio)			1.0889***	0.9933	0.9913***
			(0.0055)	(0.0045)	(0.0034)
Decomposition (as ratio) excluding Constant			1.0889	0.9076	0.9913
Male	Reference Category		1.0000	1.0005	1.0000
			(0.0000)	(0.0018)	(0.0000)
Female	-0.0066	-0.0087	1.0000	0.9994	1.0000
	(0.0053)	(0.0054)	(0.0000)	(0.0020)	(0.0000)
Sex - total effect			1.0000	0.9999	1.0000
			(0.0000)	(0.0003)	(0.0000)
Age Groups					
[0, 5)	Reference Category		1.0000	1.0048***	1.0000
			(0.0001)	(0.0010)	(0.0001)
[5, 10)	-0.0138	-0.0240	1.0000	1.0016***	1.0000
	(0.0194)	(0.0200)	(0.0001)	(0.0005)	(0.0001)
[10, 20)	0.0458***	0.0741***	1.0000	1.0065***	0.9989***
	(0.0151)	(0.0163)	(0.0001)	(0.0009)	(0.0002)
[20, 30)	-0.0091	-0.0792***	1.0003***	1.0005	1.0000
	(0.0143)	(0.0150)	(0.0001)	(0.0011)	(0.0001)
[30, 40)	0.0107	-0.0764***	1.0006***	0.9988	1.0002
	(0.0139)	(0.0150)	(0.0001)	(0.0012)	(0.0002)
[40, 50)	0.0194	-0.0811***	1.0001**	0.9970**	1.0001*
	(0.0137)	(0.0143)	(0.0001)	(0.0013)	(0.0001)
[50, 60)	0.0680***	-0.1077***	1.0002**	0.9875***	0.9992***
	(0.0137)	(0.0142)	(0.0001)	(0.0013)	(0.0002)
[60, 70)	0.1016***	-0.0456***	1.0007***	0.9897***	0.9991***
	(0.0136)	(0.0141)	(0.0001)	(0.0014)	(0.0002)
[70, 80)	0.1017***	-0.0177	1.0003***	0.9937***	0.9998**
	(0.0138)	(0.0144)	(0.0001)	(0.0015)	(0.0001)
[80, 120)	0.1406***	0.0715***	1.0012***	1.0007	1.0001
	(0.0143)	(0.0148)	(0.0002)	(0.0014)	(0.0001)
Age - total effect			1.0034***	0.9807***	0.9974***
			(0.0004)	(0.0022)	(0.0005)
Episodes					
Elective - Day Cases	0.0373***	0.0626***	1.0046***	1.0175***	1.0031***
	(0.0017)	(0.0016)	(0.0004)	(0.0016)	(0.0004)
Elective - Inpatient	0.0284***	0.1786***	0.9966***	1.0638***	0.9821***
	(0.0007)	(0.0023)	(0.0004)	(0.0025)	(0.0023)
Non-Elective - Short	-0.2610***	-0.2558***	0.9908***	1.0010	1.0002

Table continues in following page.

Table 4: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
	(0.0056)	(0.0054)	(0.0007)	(0.0014)	(0.0003)
Non-Elective - Long	0.2374***	0.2669***	1.0260***	1.0224***	1.0032***
	(0.0026)	(0.0025)	(0.0021)	(0.0028)	(0.0005)
Episodes - total effect			1.0178***	1.1068***	0.9885***
			(0.0023)	(0.0050)	(0.0024)
Bed Days					
Elective	0.0426***	0.0215***	0.9953***	0.9802***	1.0023***
	(0.0006)	(0.0004)	(0.0013)	(0.0008)	(0.0007)
Non-Elective	0.0134***	0.0092***	1.0027***	0.9831***	0.9992***
	(0.0003)	(0.0003)	(0.0010)	(0.0015)	(0.0003)
Bed Days - total effect			0.9980	0.9636***	1.0015**
			(0.0016)	(0.0016)	(0.0007)
Diagnoses					
First Diagnosis (See Appendix A for descriptions.)					
ISHMT0100	0.2573***	0.1428***	1.0051***	0.9974***	0.9978***
	(0.0181)	(0.0141)	(0.0004)	(0.0005)	(0.0005)
ISHMT0200	0.6075***	0.4056***	1.0065***	0.9821***	0.9978***
	(0.0104)	(0.0103)	(0.0010)	(0.0013)	(0.0004)
ISHMT0300	0.1171***	0.0415**	1.0004***	0.9987***	0.9997**
	(0.0200)	(0.0189)	(0.0001)	(0.0005)	(0.0001)
ISHMT0400	0.3158***	0.1352***	1.0006***	0.9968***	0.9996***
	(0.0200)	(0.0195)	(0.0002)	(0.0005)	(0.0001)
ISHMT0500	-0.1257***	0.2464***	0.9982***	1.0011***	1.0055***
	(0.0474)	(0.0217)	(0.0007)	(0.0002)	(0.0008)
ISHMT0600	0.3625***	0.3056***	1.0008**	0.9983**	0.9999*
	(0.0159)	(0.0157)	(0.0003)	(0.0007)	(0.0001)
ISHMT0700	0.1896***	0.1307***	1.0004	0.9967***	0.9999
	(0.0124)	(0.0126)	(0.0002)	(0.0010)	(0.0001)
ISHMT0800	0.3712***	0.2849***	0.9999	0.9991**	1.0000
	(0.0255)	(0.0265)	(0.0002)	(0.0004)	(0.0000)
ISHMT0900	0.5014***	0.4890***	0.9990	0.9987	1.0000
	(0.0103)	(0.0103)	(0.0008)	(0.0015)	(0.0000)
ISHMT1000	0.3429***	0.2953***	1.0045***	0.9958***	0.9994***
	(0.0108)	(0.0105)	(0.0006)	(0.0013)	(0.0002)
ISHMT1100	0.1848***	0.0587***	1.0028***	0.9772***	0.9981***
	(0.0082)	(0.0080)	(0.0004)	(0.0021)	(0.0003)
ISHMT1200	0.2714***	0.1605***	0.9994**	0.9960***	1.0002**
	(0.0144)	(0.0151)	(0.0003)	(0.0008)	(0.0001)
ISHMT1300	0.6500***	0.4941***	1.0005	0.9828***	0.9999

Table continues in following page.

Table 4: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
	(0.0096)	(0.0097)	(0.0011)	(0.0015)	(0.0003)
ISHMT1400	0.3052***	0.3172***	0.9971***	1.0012	0.9999
	(0.0096)	(0.0101)	(0.0005)	(0.0014)	(0.0001)
ISHMT1500	0.1835***	0.1772***	0.9985***	0.9998	1.0000
	(0.0165)	(0.0192)	(0.0002)	(0.0008)	(0.0002)
ISHMT1600	0.6070***	0.3503***	1.0007***	0.9989***	0.9997***
	(0.0410)	(0.0378)	(0.0002)	(0.0002)	(0.0001)
ISHMT1700	0.7685***	0.6498***	0.9998	0.9988***	1.0000
	(0.0260)	(0.0273)	(0.0004)	(0.0004)	(0.0001)
ISHMT1800	0.1224***	0.0576***	1.0005*	0.9887***	0.9997*
	(0.0082)	(0.0081)	(0.0003)	(0.0020)	(0.0001)
ISHMT1900	0.5338***	0.5466***	1.0011	1.0014	1.0000
	(0.0099)	(0.0098)	(0.0009)	(0.0016)	(0.0000)
ISHMT2100	0.1779***	0.1039***	0.9990***	0.9957***	1.0004***
	(0.0115)	(0.0123)	(0.0002)	(0.0010)	(0.0001)
First Diagnosis - total effect			1.0148***	0.9088***	0.9977**
			(0.0018)	(0.0090)	(0.0012)
Comorbidities	0.0750***	0.0776***	1.0528***	1.0032	1.0018
	(0.0023)	(0.0020)	(0.0018)	(0.0036)	(0.0021)
Type of Provider					
Large Provider	Reference Category		1.0003	0.9842***	1.0020***
			(0.0005)	(0.0044)	(0.0006)
Medium Provider	-0.0070	0.0118*	1.0003	0.9930**	1.0005*
	(0.0064)	(0.0070)	(0.0002)	(0.0035)	(0.0003)
Small Provider	0.0043	-0.0085	1.0000	0.9938***	1.0006***
	(0.0088)	(0.0096)	(0.0001)	(0.0016)	(0.0002)
Specialist Provider	0.2222***	0.2026***	1.0000	0.9982***	1.0000
	(0.0161)	(0.0169)	(0.0002)	(0.0006)	(0.0001)
Teaching Provider	0.0740***	0.0690***	1.0040***	0.9892***	0.9971***
	(0.0069)	(0.0069)	(0.0007)	(0.0028)	(0.0008)
Other Provider	-0.2530***	0.0302	0.9954***	1.0005***	1.0043***
	(0.0589)	(0.0201)	(0.0009)	(0.0001)	(0.0009)
Type of Provider - total effect			0.9999	0.9595***	1.0045***
			(0.0012)	(0.0104)	(0.0013)
Constant	6.4940***	6.6083***	-	1.0857***	-
	(0.0133)	(0.0137)		(0.0150)	
Adjusted R-squared	0.6055	0.6182			
N	66,079	75,343		141,422	

Standard errors in parenthesis. ***, **, * represent 1%, 5% and 10% significance, respectively.

The impact of type of episode on expenditures differs across the two financial years. For example, the effect of elective day case episodes on expenditure is 1.7 times greater in 2014/15 than in 2007/08; the effect of elective inpatient cases is six times greater in 2014/15 than 2007/08; elective bed days attracts half the effect on expenditure in 2014/15 than in 2007/08; and the impact of non-elective bed days in 2014/15 is 70% of the impact observed in 2007/08. Overall elective stays display a greater positive relationship with expenditure in 2014/15 than in 2007/08, while bed days (elective and non-elective) have a reduced role in 2014/15 compared to 2007/08. Non-electives have a similar relationship with expenditure in both years. Not surprisingly, the regression coefficients attached to the diagnoses dummy variables are highly statistically significant and positive. The coefficients do, however, vary across the two financial years - for example, the coefficient for ISHMT1100 (diseases of the digestive system) is 0.1848 in 2007/08 and much reduced at 0.0587 in 2014/15. The impact of comorbidities on log expenditure is greater in 2014/15 (at 0.0811) than in 2007/08 (at 0.0734). In general, the magnitudes of the effects of diagnoses outweigh the effects of other coefficients in the model. Of the variables representing provider type, only specialist, teaching and other providers display a significant relationship with expenditure. For specialist and teaching their impact on expenditure remains stable across the two periods. For other providers the effect is negative and statistically significant for expenditure in 2007/08 but small, positive and not significant in 2014/15.

The final three columns present decomposition results. The first row of Table 4 reproduces the summary decomposition results reported in Table 2 including significance levels and standard errors, and the second row shows this decomposition excluding the constant term in the results, which only affects the results for the Coefficients. The summary results show that the 7.2% increase in expenditure observed across the two periods can be decomposed into a substantial effect due to a change in the distribution of characteristics (8.9%) and a marginal contribution due to a change in the relationship between characteristics and expenditure (overall coefficient effect = -0.7%). The interactions between these two effects also contributes marginally (-0.9%). It should be noted, however, that the contribution for a change in coefficients (-0.7%) contains the contribution for a change in the constant. This effect is large at 8.5% and represents the increase in expenditure across the two periods that remains unexplained by the set of covariates included in the model. It also implies that there is a large role for a change in coefficients on the set of explanatory variables to offset the change in the constant. This effect is of magnitude -9.2% and is substantial, implying efficiency gains in expenditure across at least some of the set of variables considered. Accordingly, both the distribution of characteristics (leading to an increase in costs) and the change in coefficients (a change in the constant leading to growth in costs; a change in coefficients of covariates leading to decreased costs) contribute to the change in expenditures observed across the period.

From the detailed breakdown of the covariates we can see that the overall change in the distribution of sex and age play a negligible role in explaining the change in expenditure across the two periods. Similarly a change in the relationship between sex and expenditures (coefficient effect) is negligible. However, the impact of a change in the coefficient for age (taken across all age groups) is -2%, suggesting more efficient treatments for age in 2014/15 compared to 2007/08. This appears to be due mainly to older age groups ([50, 80)) where coefficient ratios are below 1.

A change in coefficient effects contributes significantly to the change in overall expenditure for elective patients (both day cases and inpatient care cases). The coefficient on day cases is 1.0175, hence an increase in the cost of treating day cases contributes 1.8% to the overall 7.2% growth in expenditures.

The corresponding contribution of a change in the cost of treating elective inpatients is 6.3%. The role of a change in the distribution (characteristics column) of these two variables in explaining the change in expenditure is minimal. Long stay non-elective patients contribute to the observed change in expenditures across the two periods through a change both in the distribution of characteristics and a change in the effect of these characteristics, with both explaining around 2% of the observed expenditure change. Accordingly, there appears to have been an increase in non-elective long stay patients coupled with a higher cost of treatment for such patients. Taken as a whole the impact of a change in the distribution of episodes contributes approximately 1.8% towards the change in expenditure, while a change in coefficients is substantial at 10.7%. The former is largely due to an increase in non-elective long stay patients; the latter is mainly driven by an increase in the cost of treating elective inpatient stays. The overall effect observed for episodes is partially offset by a decrease in the costs of treating elective and non-elective bed days (coefficient: 0.9802 for the former and 0.9831 for the latter: an overall contribution to the change in expenditure of -3.6%). The impact of a change in the distribution of bed days has a negligible impact on the overall change in expenditure.

The impact of a change in the distribution of diagnoses is approximately 1.5% implying a shift towards diagnoses which contribute to expenditure growth. However, the overall effect of a change in coefficients (the relationship between diagnoses and expenditure) is approximately -9.1%, implying greater efficiency in treating conditions. Notable contributions to the decomposition analysis are observed for ISHMT0200, ISHMT1100, ISHMT1300 and ISHMT1800 (respectively, neoplasms; diseases of the digestive system; musculoskeletal system and connective tissue; symptoms, signs and abnormal clinical and laboratory findings not elsewhere classified).[15] For all of these diagnoses a change in the coefficient (the relationship between diagnosis and expenditure) is negative and statistically significant. However, the change in the distribution of these diagnoses is positive. Accordingly, while it appears there is a small increase in the incidence of these diagnoses, this is outweighed by a decrease in the costs of treatment.

A major contributor to the decomposition analysis is the recorded number of comorbidities. While the contribution of a change in coefficients (relationship with expenditure) is small (1.0032), the change in characteristics (the distribution of the number of comorbidities) is large (1.0528) and significant contributing 5.3% to the 7.2% rise in expenditure across the period. This results from an increase in the average number of reported comorbidities across the two periods. A comparison of models (3) and (4) in Table 2 provides further insight into the effects of comorbidities. The two models are identical except that model (4) includes comorbidities. A decomposition of the 7.2% increase in expenditure across the two periods in model (3) reveals that 4.3% of this is due to a change in characteristics and 2.3% is due to a change in coefficients (the remainder (0.6%) is due to the interaction of these two effects). A change in characteristics dominates the change in coefficients. However, when the number of comorbidities is included, the change in characteristics contributes 11.5% to the overall change in expenditure, with a change in coefficients contributing negatively (-3.5%). That is, the inclusion of comorbidities places far greater weight on the impact of characteristics, with coefficients contributing to a lowering of expenditure across the two periods.

A change in the distribution of patients attending different providers has a negligible effect on the observed change in expenditure. However, the change in coefficients for providers contributes to lowering costs across the period by approximately 4.1%. This effect is a combination of a decrease in costs of treating patients over time in all providers except those designated Other Providers.

[15] Notable implying an absolute contribution of greater than 1% to the overall observed change in expenditure across the two periods.

While the overall summary of the decompositions for model (7) suggests a substantial role for the change in the distribution of characteristics in explaining the observed change in expenditure, when considering the detailed results for the individual variables, there are few estimated effects that appear meaningful. The exceptions are the effects of a change in the distribution of comorbidities which contributes approximately 5% to the overall result and long stay non-elective patients who contribute approximately 2.6% to the overall change in expenditures. For both these characteristics, the change in coefficients is also positive (albeit small for comorbidities). This indicates that these effects have driven the rise in expenditures through both a characteristic and coefficient effect. However, of the 41 covariates included in the model, 31 have estimated ratios for characteristics that are equal to or greater than unity. Accordingly, for the majority of covariates changes in their distribution over time contribute to the growth in expenditure, even if only marginally. Conversely, for 29 of the 44 covariates the change in coefficients implies a reduction in expenditures over time. The major contributory factor to the growth in expenditure among the coefficient effects is attached to the change in the constant term which in itself contributes 8.6%. This effect offsets the reduction in expenditure due to the coefficient effects estimated for the individual variables (-9.2%), such that the overall coefficient effect contributes less than 1% of the overall 7.2% growth in expenditures. Accordingly, while many covariates indicate more efficient delivery of health care over time, this is offset by the unexplained change in the constant.

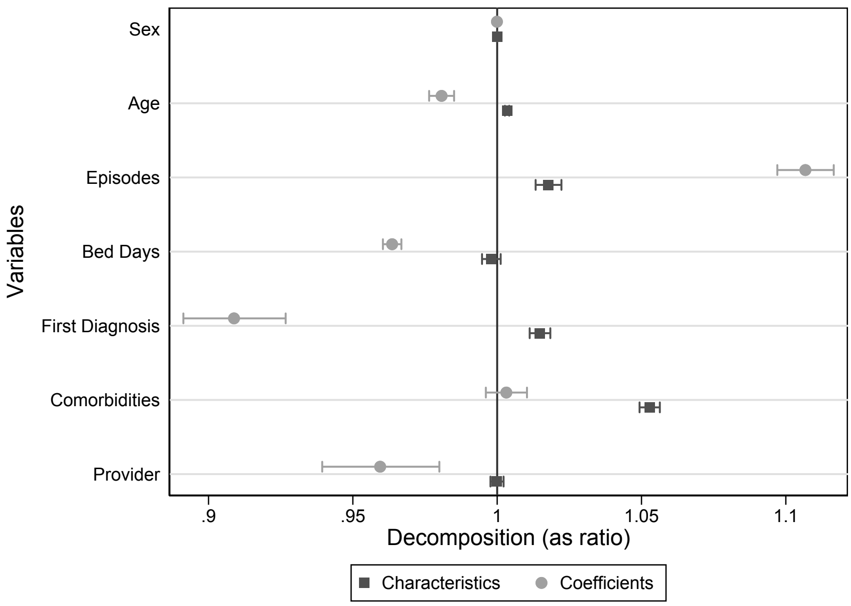


Figure 3: Summary of Table 4

Figure 3 summarizes the results shown in Table 4. For each group of variables it shows the estimate (with confidence interval) of the decomposition into Characteristics and Coefficients. The vertical line at 1 indicates no effect, i.e. estimates which confidence interval cross this line are not significant.

6.2. Decomposition of the full distribution of expenditures

Table 5 shows the results based on the decomposition method of Chernozhukov et al. (2013) using the same data and specification (model 7) as for the Oaxaca Decomposition.[16] The approach allows the

[16] We implement this approach using the Stata command *cdeco*.

change in expenditures to be decomposed across the full distribution, rather than focusing on the change in expenditures at the mean. This is achieved by estimating the counterfactual distribution using the conditional distribution, $F_{Y(y|x)}$, in one group and the distribution of independent variables $(F_{X(x)})$ of the other group. For the two periods of HES data, the decomposition is set out in Equation (4), where the index 1 denotes the year 2014/15 and 0 the year 2007/08. To simplify presentation, Table 5 reports the difference in observed outcomes and the decomposition at the deciles of the distribution. These are calculated on a log scale. A more detailed breakdown of the results estimated at each percentile is illustrated in Figure 4.

Table 5: Counterfactual Decomposition of Differences in Distributions - Log(Hospital Costs) between 2007/08 and 2014/15

Quantile	Differences between the observable distributions	Effects of characteristics	Effects of coefficients
0.1	-0.106	-0.004	-0.102
	[-0.119, -0.093]	[-0.010, 0.002]	[-0.113, -0.090]
0.2	-0.045	0.002	-0.047
	[-0.056, -0.033]	[-0.004, 0.008]	[-0.057, -0.036]
0.3	0.014	0.013	0.000
	[0.002, 0.025]	[0.007, 0.019]	[-0.009, 0.010]
0.4	0.056	0.022	0.035
	[0.044, 0.069]	[0.015, 0.028]	[0.026, 0.044]
0.5	0.085	0.026	0.060
	[0.073, 0.098]	[0.019, 0.032]	[0.051, 0.069]
0.6	0.104	0.030	0.074
	[0.091, 0.117]	[0.023, 0.038]	[0.065, 0.082]
0.7	0.121	0.042	0.078
	[0.106, 0.135]	[0.032, 0.052]	[0.070, 0.087]
0.8	0.146	0.078	0.068
	[0.124, 0.168]	[0.064, 0.093]	[0.053, 0.083]
0.9	0.211	0.177	0.034
	[0.172, 0.250]	[0.152, 0.203]	[-0.001, 0.068]
N 2007/08		66,079	
N 2014/15		75,343	

Pointwise Confidence Interval in parenthesis.

The first two quantiles of Table 5 show an overall negative change in expenditures (a decrease in expenditure from 2007/08 to 2014/15). The change is then positive and increasing in magnitude thereafter (an observed growth in expenditure). Larger increases are observed at the top of the expenditure distribution. That is, expensive to treat patients have raised expenditure disproportionately more than less expensive patients. All of the observed effects are statistically significant (none of the confidence intervals straddle zero). The final two columns show the decomposition results for the effect of a change in characteristics and the effect for a change in coefficients. Again, from the 30th quantile upwards, these effects are positive and statistically significant implying their contribution increases expenditures. In general the changes in expenditure across the quantiles are driven by a change in coefficients rather than a change in characteristics (coefficient effects have higher absolute values than the effect of characteristics).[17] For high expenditure patients (80th and 90th quantile) the effect of a change in characteristics is more important in explaining the rise in expenditure than the effect for a change in coefficients. This is particularly evident at the 90th quantile. The increase in expenditure at the top of the distribution is driven largely by changes in the composition of patients. At the bottom of the distribution coefficient effects dominate the effects of characteristics implying that expenditure savings over the two periods are driven by more efficient treatment of patients. In the middle of the distribution, again the effect of coefficients dominates characteristics. However, here both effects contribute to increasing, rather than decreasing expenditures.

To better understand the decomposition results across the full distribution, Table 6 summarises the differences in the key covariates of the model across the two time periods broken down into the bottom 10%, the middle (10th to 90th percentile), and the top 10% of the expenditure distribution in each year. This illustrates the large increase in expenditure at the top end of the distribution (4,142.6) compared to other parts of the distribution (-41.9 in the bottom 10% and 289.3 in the middle of the distribution). While the average age of patients treated decreased in the lower part of the expenditure distribution, they increased more substantially in the middle compared to the top of the distribution. The number of comorbidities and the number of first diagnoses increase across the two periods to a greater extent at the top of the expenditure distribution than in other areas of the distribution. This emphasises the role of comorbidity in driving expenditure increases. Other notable increases that are particularly pronounced at the top of the distribution are the number of episodes per patient, particularly the average number of non-elective episodes.[18] The number of non-elective episodes per patient increased on average by 0.92, while the average number of non-elective bed days per patient increased by 2.47 across the period in the top 10% of the expenditure distribution. Compared to the middle of the distribution (where for example, non-elective bed days decrease), these increases are notable.

Figure 4 plots the overall change in expenditures together with the effects of coefficients and characteristics across the full expenditure distribution. For the majority of the distribution (to the 80th quantile), the plotted lines for the observed change in expenditures and the coefficients effects overlap emphasising the strong contribution of coefficients to the overall change. In contrast, the plot for the effects of a change in characteristics across the quintile distribution is close to zero, except at the top of the distribution (80th quantile onwards) where they closely align to the plot for the overall change in expenditures.

[17] Note that for model 7, the Oaxaca decomposition at the mean showed a larger role for the effects of a change in characteristics rather than a change in coefficients.

[18] Note that the average number of elective episodes decreases between 2007/08 and 2014/15 at the top of the distribution, due to a shift from inpatient to daycase care.

Table 6: Differences in Mean between 2007/08 and 2014/15

	Bottom 10%		Middle [10th - 90th pctile]		Top 10%	
	Diff.Mean	StdErr	Diff.Mean	StdErr	Diff.Mean	StdErr
Cost (in £ of 2007/08)	-41.86	0.93	289.26	10.37	4142.57	199.77
Demographics						
Age	-3.56	0.40	2.33	0.15	1.58	0.36
Male	-0.0309	0.0084	0.0042	0.0030	-0.0084	0.0084
Morbidity						
Num. different 1st diagnoses						
based in 20 ISHMT chapters	0.00	0.00	0.05	0.00	0.25	0.02
based in 130 ISHMT groups	0.00	0.00	0.06	0.00	0.30	0.02
Number of Comorbidities *	0.27	0.02	0.66	0.01	1.35	0.04
Activity						
Total Episodes	0.00	0.00	0.15	0.01	0.25	0.17
Elective Episodes	-0.01	0.01	0.09	0.01	-0.66	0.17
Daycases	-0.02	0.01	0.11	0.01	0.36	0.08
Inpatient	0.00	0.00	-0.02	0.00	-1.02	0.15
Non-Elective Episodes	0.02	0.01	0.06	0.01	0.92	0.06
Short Stay	0.02	0.01	0.03	0.00	0.06	0.01
Long Stay	0.00	0.00	0.03	0.01	0.85	0.05
Bed Days	0.07	0.05	-0.21	0.03	2.47	0.58
Elective	0.03	0.01	-0.14	0.01	0.00	0.28
Non-Elective	0.05	0.05	-0.06	0.03	2.47	0.54
Provider Type						
Large Provider	-0.0373	0.0081	-0.0469	0.0028	-0.0472	0.0079
Medium Provider	-0.0247	0.0075	-0.0203	0.0026	-0.0196	0.0072
Small Provider	-0.0014	0.0053	-0.0121	0.0018	-0.0063	0.0050
Specialist Provider	0.0016	0.0017	-0.0001	0.0010	-0.0029	0.0031
Teaching Provider	0.0473	0.0072	0.0618	0.0026	0.0526	0.0076
Other Provider	0.0144	0.0016	0.0176	0.0006	0.0233	0.0020
Num. Obs. 2007/08	6,607		52,864		6,608	
Num. Obs. 2014/15	7,534		60,274		7,535	

*based on 130 ISHMT groups.

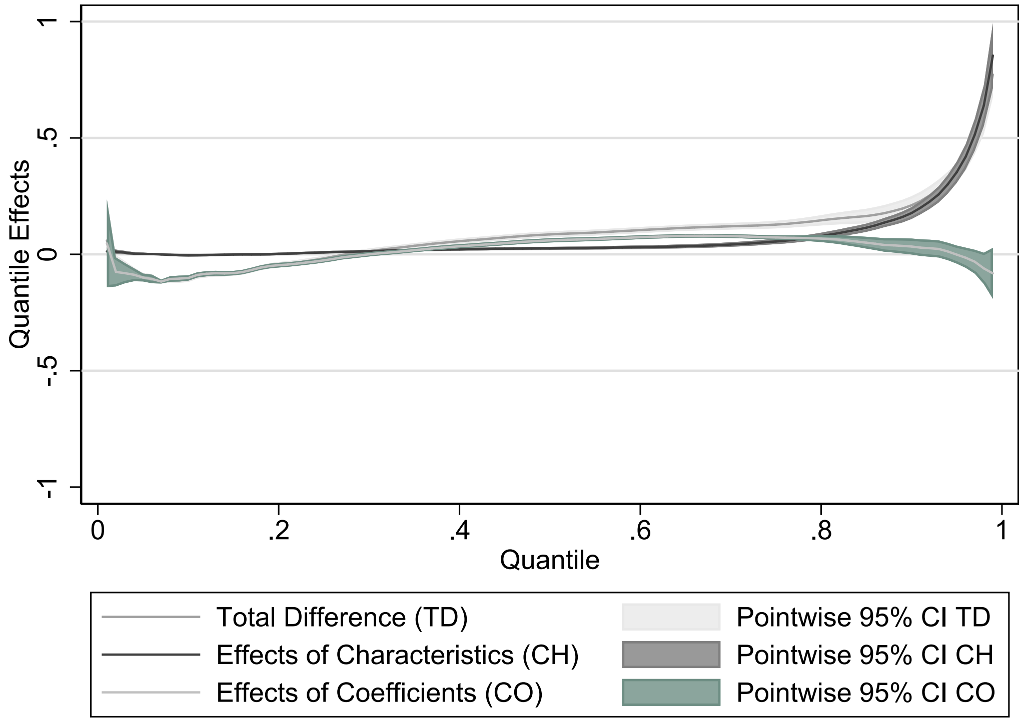


Figure 4: Counterfactual Decomposition of Differences in Distributions - Log(Hospital Costs) between 2007/08 and 2014/15

7. Discussion

This paper considers the detailed breakdown of hospital inpatient expenditures across the period 2007/08 to 2014/15. Decomposition techniques are used to unpick the observed growth in expenditure into components due to a change in the distribution of characteristics (for example, greater prevalence of morbidity measured via diagnostic categories) and due to a change in the impact of such characteristics on expenditures (coefficient effects). We undertake this analysis by randomly sampling 1% of the population of hospital inpatient users in each of the two years. The overall growth in expenditure per patient from 2007/08 to 2014/15 was £641.60, representing a 21.6% (7.23% on the log scale) increase over the period.

For the most saturated model (model (7); Table 4) we observe statistically significant relationships with expenditure due to age, type of episode (elective - day cases; elective - inpatient; non-elective - short; non-elective - long), bed days, diagnoses, the number of comorbidities, and certain types of provider (specialist, teaching). The majority of the characteristics are specified as dummy variables and, hence, coefficients can be readily compared. In general, the magnitude of effects for diagnoses outweighs effects for other coefficients in the model in explaining expenditure.

The decomposition analysis shows that the increase in (log) expenditure observed across the two periods can be decomposed into a substantial effect due to a change in the distribution of characteristics and a marginal contribution due to a change in the relationship between characteristics and expenditure. It should be noted, however, that the contribution of a change in coefficients contains the contribution of a change in the constant term. This effect is large, almost as large as the effect due to a change in the distribution of characteristics, and represents the increase in expenditure across the two periods that remains unexplained by the set of covariates included in the model. It also implies that there is a large role for a change in coefficients on the set of explanatory variables to offset the change in the constant, implying efficiency gains in expenditure across at least some of the set of variables considered. Accordingly, both the distribution of characteristics (leading to an increase in costs) and the change in coefficients (constant leading to an increase in costs; covariates to a decrease in costs) contribute to the change in expenditures observed across the period.

Overall, sex and age play a negligible role (both in a change in characteristics and change in coefficients) in explaining the increase in expenditure across the two periods. Expenditure on elective inpatient care (both day cases and inpatient care cases) has risen, mainly due to an increase in treatment costs rather than a structural change in a shift towards elective treatments. Similarly expenditure on non-elective long stay episodes has increased, due to both an increase in such cases together with an increase in the relationship with expenditure. The overall impact of a change in the distribution of diagnoses implies a modest shift towards diagnoses which contribute to increased expenditure. The corresponding overall effect of a change in coefficients (the relationship between diagnoses and expenditure), on the other hand, implies greater efficiency in treating conditions. This is particularly the case for neoplasms (ISHMT0200), diseases of the digestive system (ISHMT1100), musculoskeletal system and connective tissue (ISHMT1300) and symptoms, signs and abnormal clinical and laboratory findings not elsewhere classified (ISHMT1800).

A major contributor to the decomposition analysis is the recorded number of comorbidities. The effect of a structural change in the distribution of the number of comorbidities is large and significant, contributing more than half of the overall rise in expenditure due to all characteristics across the period. This results from an increase in the average number of reported comorbidities. Aragón et al. (2016) document the rise in the average number of comorbidities reported in Hospital Episode Statistics over the period 1998/99 to 2012/13. They note the consistent increase in the number of reported comorbidities over the period from around an average of 0.2 for women and 0.3 for men in 1998/99 to approximately 0.3 and 0.4 respectively at the end of the series. They also document the increased expenditures associated with comorbidities - in 2012/13 the difference in average expenditure for patients recorded without a comorbidity and with a single comorbidity was £1,790; the difference in average expenditure for patients with 1 and 2 comorbidities was £3,478; between 2 and 3, £4,096; between 3 and 4, £4,677 and between 4 and 4+ £6,133. Clearly comorbidities are more complex and expensive to treat. The observed increase in comorbidity may help to explain the observed rise in activity (episodes of care). However, it is important to note that the recording of diagnoses changed between the two years considered in this analysis, which likely led to an increase in the number of comorbidities documented in HES (refer to footnote 7).

While the overall summary of the decompositions for the most detailed model (model (7)) suggests a substantial role for a structural change in the distribution of characteristics in explaining the observed change in expenditure, the detailed results sugest few individual characteristics that appear meaningful. The exceptions are the effects of a change in the distribution of comorbidities which contributes approximately 5.3% and long stay non-elective patients contributing approximately 2.6% to the overall change in expenditures. For both these characteristics, the change in the relationship with expenditure is also positive (albeit small for comorbidities). This indicates that these effects have driven the rise in expenditures through both a structural characteristic and coefficient effect.

When considering the full distribution of expenditures, we observe a negative change at the bottom of the distribution (bottom two quintiles) but increasing expenditure thereafter. Large increases are observed at the top of the expenditure distribution. Here, the effect of changes in characteristics are more important in explaining the rise in expenditures than the effects for a change in coefficients. These changes appear due to increases in comorbidities (and the average number of first diagnoses) across the two periods, together with increases in non-elective long stay episodes and non-elective bed days.

The results of this analysis needs to be placed within the context of broader economic constraints within which the NHS operates. Decisions to access and consume health care resources is not simply based on the needs of patients. Decisions about the allocation of resources across disease groups and individuals are implicitly made within a budget constraint. Accordingly the level of overall sickness or ill-health within the population is not the only major determinant of aggregate health care expenditure. Population ageing, increased duration of time spent with chronic conditions, and multi-morbidity will all lead to shifts in the way health care expenditure is determined and allocated, but so will political preferences and economic growth which influence funding levels and the organisation of health care provision (see for example, Getzen (2001)). While comparing the determinants of expenditure across periods is not straightforward in any dynamic system, the decomposition analysis employed helps elucidate effects due to changes in the distribution of drivers and changes in how a given characteristic, such as given diagnosis, is treated. The techniques further allow investigation of where in the distribution of expenditures these effects are most prominent and where expenditure rises have been most notable.

References

Aragón, M. J., M. Chalkley and N. Rice (2016). 'Medical Spending and Hospital Inpatient Care in England: An Analysis over Time'. *Fiscal Studies*, 37: 405–432. ISSN: 1475-5890. DOI: 10.1111/j.1475-5890.2016.12102. URL: http://dx.doi.org/10.1111/j.1475-5890.2016.12102.

Blinder, A. S. (1973). 'Wage Discrimination: Reduced Form and Structural Estimates'. *The Journal of Human Resources*, 8: 436–455. ISSN: 0022166X. URL: http://www.jstor.org/stable/144855.

Charlesworth, A., L. Hawkins and L. Marshall (2014). *NHS payment reform: lessons from the past and directions for the future*. Research report. Nuffield Trust. URL: https://www.nuffieldtrust.org.uk/research/nhs-payment-reform-lessons-from-the-past-and-directions-for-the-future.

Chernozhukov, V., I. Fernandez-Val and B. Melly (2012). *Inference on Counterfactual Distributions*. URL: http://www.mit.edu/~vchern/papers/counterfactual_2012Nov1.pdf.

Chernozhukov, V., I. Fernández-Val and B. Melly (2013). 'Inference on Counterfactual Distributions'. *Econometrica*, 81: Stata commands: https://sites.google.com/site/blaisemelly/computer-programs/inference-on-counterfactual-distributions, 2205–2268. ISSN: 1468-0262. DOI: 10.3982/ECTA10582. URL: http://dx.doi.org/10.3982/ECTA10582.

de Meijer, C., O. O'Donnell, M. Koopmanschap and E. van Doorslaer (2013). 'Health expenditure growth: Looking beyond the average through decomposition of the full distribution'. *Journal of Health Economics*, 32: 88–105. ISSN: 0167-6296. DOI: http://dx.doi.org/10.1016/j.jhealeco.2012.10.009. URL: http://www.sciencedirect.com/science/article/pii/S0167629612001701.

Department of Health (2008). *NHS reference costs 2006-07*. URL: http://webarchive.nationalarchives.gov.uk/20130104223439/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_082571.

— (2009). *NHS reference costs 2007-08*. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_098945.

— (2014). *NHS reference costs*. URL: https://www.gov.uk/government/collections/nhs-reference-costs.

— (2015). *NHS reference costs 2014 to 2015*. URL: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015.

— (2016). *HCHS Pay & Price Inflation*. URL: http://www.info.doh.gov.uk/doh/finman.nsf/af3d43e36a4c8f8500256722005b77f8/360a47827991d10a80258036002d8d9f/$FILE/2015.16%20Pay%20&%20Price%20series.xlsx.

DH PbR Team (2009). *Payment by Results Guidance 2009-10*. URL: http://webarchive.nationalarchives.gov.uk/+/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_094183.

Dormont, B., M. Grignon and H. Huber (2006). 'Health expenditure growth: reassessing the threat of ageing'. *Health Economics*, 15: 947–963. ISSN: 1099-1050. DOI: 10.1002/hec.1165. URL: http://dx.doi.org/10.1002/hec.1165.

Foresi, S. and F. Peracchi (1995). 'The Conditional Distribution of Excess Returns: An Empirical Analysis'. *Journal of the American Statistical Association*, 90: 451–466. ISSN: 01621459. URL: http://www.jstor.org/stable/2291056.

Fortin, N., T. Lemieux and S. Firpo (2011). 'Chapter 1 - Decomposition Methods in Economics'. In: ed. by O. Ashenfelter and D. Card. Vol. 4, Part A. Handbook of Labor Economics. Elsevier, 1–102. DOI: https://doi.org/10.1016/S0169-7218(11)00407-2. URL: http://www.sciencedirect.com/science/article/pii/S0169721811004072.

Getzen, T. E. (2001). 'Aging and health care expenditures: A comment on Zweifel, Felder and Meiers'. *Health Economics*, 10: 175–177. ISSN: 1099-1050. DOI: 10.1002/hec.593. URL: http://dx.doi.org/10.1002/hec.593.

Health and Social Care Information Centre (2017). *Coding Clinic*. Tech. rep. URL: https://hscic.kahootz.com/connect.ti/t_c_home/view?objectId=297843#297843.

HM Treasury (2010-2017). *Public Expenditure Statistical Analyses (PESA)*. URL: https://www.gov.uk/government/collections/public-expenditure-statistical-analyses-pesa.

Howdon, D. and N. Rice (2017). 'Health care expenditures, age, proximity to death and morbidity: Implications for an ageing population'. *Journal of Health Economics*, ISSN: 0167-6296. DOI: https://doi.org/10.1016/j.jhealeco.2017.11.001. URL: http://www.sciencedirect.com/science/article/pii/S0167629617310020.

Jann, B. (2008). 'The Blinder-Oaxaca decomposition for linear regression models'. *Stata Journal*, 8: 453–479(27). URL: http://www.stata-journal.com/article.html?article=st0151.

Licchetta, M. and M. Stelmach (2016). *Fiscal sustainability and public spending on health*. Fiscal sustainability analytical paper. Office for Budget Responsibility. URL: http://budgetresponsibility.org.uk/docs/dlm_uploads/Health-FSAP.pdf.

Monitor and NHS England (2013). *National tariff payment system 2014/15*. URL: http://webarchive.nationalarchives.gov.uk/20141209104711/https://www.gov.uk/government/publications/national-tariff-payment-system-2014-to-2015.

National Casemix Office (2015). *Casemix Companion*. URL: http://content.digital.nhs.uk/article/6226/HRG4-201415-Reference-Cost-Grouper.

National Casemix Office (2015). *HRG4+ 2014/15 Reference Cost Grouper*. URL: http://content.digital.nhs.uk/article/6226/HRG4-201415-Reference-Cost-Grouper.

NHS Digital (2015). *Hospital Episode Statistics, Admitted Patient Care - England, 2014-15*. URL: http://content.digital.nhs.uk/catalogue/PUB19124/hosp-epis-stat-admi-summ-rep-2014-15-rep.pdf.

NHS England (n.d.[a]). *Bed Availability and Occupancy Data - Day only*. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-day-only/.

— (n.d.[b]). *Bed Availability and Occupancy Data - Overnight*. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/.

Oaxaca, R. (1973). 'Male-Female Wage Differentials in Urban Labor Markets'. *International Economic Review*, 14: 693–709. ISSN: 00206598, 14682354. URL: http://www.jstor.org/stable/2525981.

Oaxaca, R. L. and M. R. Ransom (1999). 'Identification in Detailed Wage Decompositions'. *The Review of Economics and Statistics*, 81: 154–157. DOI: 10.1162/003465399767923908. eprint: http://dx.doi.org/10.1162/003465399767923908. URL: http://dx.doi.org/10.1162/003465399767923908.

OECD (2006). *Projecting OECD Health and Long-Term Care Expenditures*. URL: http://www.oecd-ilibrary.org/economics/projecting-oecd-health-and-long-term-care-expenditures_736341548748.

— (2017). *OECD Health Statistics 2017: Frequently Requested Data*. URL: http://www.oecd.org/health/health-statistics.htm%20[go%20to%20OECD%20Health%20Statistics%202017:%20Frequently%20Requested%20Data].

Office for National Statistics (2013). *Revised Annual Mid-year Population Estimates: 2001 to 2010*. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/2013-12-17.

— (2015). *Annual Mid-year Population Estimates: 2014*. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/2015-06-25.

— (2017). *Volume growth, contributions to growth and expenditure shares for public service healthcare quantity output by component*. URL: https://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity/datasets/volumegrowthcontributionstogrowthandexpendituresharesforpublicservicehealthcarequantityoutputbycomponent.

StataCorp. (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

The Casemix Service (2008). *HRG4 2007/08 Reference Costs Grouper*. URL: http://content.digital.nhs.uk/article/2322/HRG4-200708-Reference-Costs-Grouper-Documentation%20http://content.digital.nhs.uk/article/5991/HRG4-200708-Reference-Costs-Grouper.

The Health and Social Care Information Centre (2009). *Headline figures, 2007-08*. URL: http://content.digital.nhs.uk/catalogue/PUB02541/hosp-epis-stat-admi-head-figs-07-08-rep.pdf.

World Health Organization (n.d.). *International Shortlist for Hospital Morbidity Tabulation (ISHMT)*. URL: http://www.who.int/classifications/icd/implementation/morbidity/ishmt/en/.

Yun, M.-S. (2005). 'A Simple Solution to the Identification Problem in Detailed Wage Decompositions'. *Economic Inquiry*, 43: 766–772. ISSN: 1465-7295. DOI: 10.1093/ei/cbi053. URL: http://dx.doi.org/10.1093/ei/cbi053.

A. Appendix A

Table A.1 shows the descriptions of the 20 ISHMT chapters used to classify diagnoses in the main part. For a description of the groups within each chapter see `http://www.who.int/classifications/icd/implementation/morbidity/ishmt/en/`.

Table A.1: ISHMT Classification

ISHMT Code	Heading	ICD Codes
0100	Certain infectious and parasitic diseases	A00 - B99
0200	Neoplasms	C00 - D48
0300	Diseases of the blood and bloodforming organs and certain disorders involving the immune mechanism	D50 - D89
0400	Endocrine, nutritional and metabolic diseases	E00 - E90
0500	Mental and behavioural disorders	F00 - F99
0600	Diseases of the nervous system	G00 - G99
0700	Diseases of the eye and adnexa	H00 - H59
0800	Diseases of the ear and mastoid process	H60 - H95
0900	Diseases of the circulatory system	I00 - I99
1000	Diseases of the respiratory system	J00 - J99
1100	Diseases of the digestive system	K00 - K93
1200	Diseases of the skin and subcutaneous tissue	L00 - L99
1300	Diseases of the musculoskeletal system and connective tissue	M00 - M99
1400	Diseases of the genitourinary system	N00 - N16
1500	Pregnancy, childbirth and the puerperium	O00 - O99
1600	Certain conditions originating in the perinatal period	P00 - P96
1700	Congenital malformations, deformations and chromosomal abnormalities	Q00 - Q99
1800	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	R00 - R99
1900	Injury, poisoning and certain other consequences of external causes	S00 - T98
2100	Factors influencing health status and contact with health services	Z00 - Z99

B. Appendix B

This section presents results using the 130 diagnosis groups in ISHMT rather than the 20 chapters used in Table 4 (see Section B.1) and Table 5 and Figure 4 (see Section B.2).

B.1. Oaxaca

Table B.1: Blinder-Oaxaca Decomposition of Hospital Costs Difference between 2007/08 and 2014/15 - Reference Group: 2007/08 - Specification (7) - 130 diagnosis groups

			Decomposition (as ratio)		
	2007/08	2014/15	Characteristics	Coefficients	Interaction
Decomposition (as ratio)			1.0916***	0.9939	0.9883***
			(0.0058)	(0.0041)	(0.0033)
Decomposition (as ratio) excluding Constant			1.0916	0.0584	0.9883
Male	Reference Category		1.0000	1.0012	1.0000
			(0.0000)	(0.0017)	(0.0000)
Female	-0.0205***	-0.0255***	1.0000	0.9986	1.0000
	(0.0051)	(0.0052)	(0.0000)	(0.0020)	(0.0000)
Sex - total effect			1.0000	0.9998	1.0000
			(0.0001)	(0.0002)	(0.0000)
Age Groups					
[0, 5)	Reference Category		1.0000	1.0046***	1.0000
			(0.0000)	(0.0010)	(0.0001)
[5, 10)	-0.0315*	-0.0401**	1.0000	1.0016***	1.0000
	(0.0181)	(0.0190)	(0.0000)	(0.0005)	(0.0001)
[10, 20)	0.0465***	0.0619***	0.9997***	1.0055***	0.9991***
	(0.0144)	(0.0157)	(0.0001)	(0.0009)	(0.0002)
[20, 30)	-0.0300**	-0.1150***	1.0003***	0.9988	1.0001
	(0.0137)	(0.0146)	(0.0001)	(0.0011)	(0.0001)
[30, 40)	-0.0182	-0.1149***	1.0006***	0.9974**	1.0004**
	(0.0134)	(0.0146)	(0.0001)	(0.0011)	(0.0002)
[40, 50)	-0.0098	-0.1094***	1.0001**	0.9967***	1.0001*
	(0.0132)	(0.0140)	(0.0001)	(0.0012)	(0.0001)
[50, 60)	0.0186	-0.1413***	1.0000	0.9891***	0.9993***
	(0.0132)	(0.0139)	(0.0001)	(0.0012)	(0.0002)
[60, 70)	0.0419***	-0.0883***	1.0003***	0.9917***	0.9993***
	(0.0131)	(0.0138)	(0.0001)	(0.0013)	(0.0002)
[70, 80)	0.0503***	-0.0537***	1.0002**	0.9954***	0.9998**
	(0.0133)	(0.0142)	(0.0001)	(0.0014)	(0.0001)
[80, 120)	0.1105***	0.0583***	1.0012***	1.0024*	1.0003*
	(0.0.138)	(0.0146)	(0.0002)	(0.0013)	(0.0001)
Age - total effect			1.0024***	0.9833***	0.9984***

Table continues in following page.

Table B.1: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
			(0.0003)	(0.0022)	(0.0005)
Episodes					
Elective - Day Cases	0.0347***	0.0559***	1.0043***	1.0146***	1.0026***
	(0.0016)	(0.0016)	(0.0004)	(0.0015)	(0.0003)
Elective - Inpatient	0.0278***	0.1457***	0.9967***	1.0491***	0.9859***
	(0.0007)	(0.0022)	(0.0004)	(0.0021)	(0.0018)
Non-Elective - Short	-0.2418***	-0.2417***	0.9915***	1.0000	1.0000
	(0.0054)	(0.0053)	(0.0007)	(0.0014)	(0.0003)
Non-Elective - Long	0.2231***	0.2507***	1.0244***	1.0209***	1.0030***
	(0.0026)	(0.0025)	(0.0020)	(0.0028)	(0.0005)
Episodes - total effect			1.0167***	1.0868***	0.9915***
			(0.0021)	(0.0047)	(0.0019)
Bed Days					
Elective	0.0341***	0.0192***	0.9963***	0.9860***	1.0016***
	(0.0005)	(0.0004)	(0.0010)	(0.0007)	(0.0005)
Non-Elective	0.0113***	0.0076***	1.0023***	0.9847***	0.9992***
	(0.0003)	(0.0002)	(0.0008)	(0.0014)	(0.0003)
Bed Days - total effect			0.9985	0.9709***	1.0009
			(0.0013)	(0.0016)	(0.0006)
Diagnoses					
First Diagnosis					
ISHMT1	0.2201***	0.1502***	0.9999	0.9995	1.0000
	(0.0295)	(0.0317)	(0.0001)	(0.0003)	(0.0000)
ISHMT2	0.2856***	0.0088	1.0039***	0.9997***	0.9962***
	(0.0750)	(0.0209)	(0.0010)	(0.0001)	(0.0011)
ISHMT3	0.8457***	0.6799***	0.9996***	0.9999	1.0001
	(0.0862)	(0.1442)	(0.0001)	(0.0001)	(0.0001)
ISHMT4	0.1723***	0.0812***	1.0006***	0.9997*	0.9997*
	(0.0397)	(0.0297)	(0.0002)	(0.0002)	(0.0002)
ISHMT5	0.7497***	(omitted)	0.9997***	0.9997***	1.0003***
	(0.1112)		(0.0001)	(0.0001)	(0.0001)
ISHMT6	0.2714***	0.2233***	1.0013***	0.9995	0.9998
	(0.0235)	(0.0211)	(0.0002)	(0.0003)	(0.0002)
ISHMT7	0.9466***	0.7484***	0.9998	0.9988***	1.0000
	(0.0311)	(0.0327)	(0.0004)	(0.0003)	(0.0001)
ISHMT8	0.7879***	0.5171***	0.9998	0.9986***	1.0001
	(0.0332)	(0.0353)	(0.0003)	(0.0003)	(0.0001)
ISHMT9	0.1728***	0.0488**	1.0007***	0.9987***	0.9995***
	(0.0234)	(0.0210)	(0.0001)	(0.0003)	(0.0001)

Table continues in following page.

Table B.1: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
ISHMT10	1.1246***	0.9448***	0.9999	0.9987***	1.0000
	(0.0281)	(0.0300)	(0.0005)	(0.0003)	(0.0001)
ISHMT11	0.8989***	0.9344***	1.0000	1.0001	1.0000
	(0.0623)	(0.0642)	(0.0002)	(0.0001)	(0.0000)
ISHMT12	1.0109***	0.6824***	1.0001	0.9996***	1.0000
	(0.0652)	(0.0663)	(0.0002)	(0.0001)	(0.0001)
ISHMT13	0.5941***	0.4313***	1.0005**	0.9993***	0.9999*
	(0.0364)	(0.0348)	(0.0002)	(0.0002)	(0.0001)
ISHMT14	0.5501***	0.4868***	0.9992***	0.9996	1.0001
	(0.0320)	(0.0386)	(0.0002)	(0.0003)	(0.0001)
ISHMT15	0.8503***	0.6684***	1.0026***	0.9959***	0.9994***
	(0.0172)	(0.0174)	(0.0007)	(0.0006)	(0.0002)
ISHMT16	0.4025***	0.2634***	1.0002	0.9996**	0.9999
	(0.0454)	(0.0437)	(0.0001)	(0.0002)	(0.0001)
ISHMT17	0.0183	-0.1115***	1.0001	0.9994***	0.9991***
	(0.0350)	(0.0236)	(0.0002)	(0.0002)	(0.0003)
ISHMT18	0.7179***	0.7119***	0.9996*	1.0000	1.0000
	(0.0388)	(0.0436)	(0.0002)	(0.0002)	(0.0000)
ISHMT19	0.3214***	0.2329***	0.9998	0.9983***	1.0001
	(0.0175)	(0.0186)	(0.0002)	(0.0005)	(0.0001)
ISHMT20	0.1067***	0.0293	1.0004***	0.9989***	0.9997**
	(0.0208)	(0.0192)	(0.0001)	(0.0004)	(0.0001)
ISHMT21	0.2798***	0.2568***	0.9999	0.9999	1.0000
	(0.0371)	(0.0399)	(0.0001)	(0.0002)	(0.0000)
ISHMT22	0.2831***	0.2441***	0.9996***	0.9998	1.0001
	(0.0305)	(0.0364)	(0.0001)	(0.0003)	(0.0001)
ISHMT23	0.3979***	0.1637***	1.0014***	0.9972***	0.9992***
	(0.0223)	(0.0205)	(0.0003)	(0.0004)	(0.0002)
ISHMT24	-0.6625***	0.2312***	0.9989***	1.0003***	1.0015***
	(0.1272)	(0.0556)	(0.0002)	(0.0001)	(0.0003)
ISHMT25	-0.3972***	0.3066***	0.9987***	1.0009***	1.0024***
	(0.0670)	(0.0373)	(0.0002)	(0.0001)	(0.0003)
ISHMT26	1.3321***	0.2967***	1.0011***	1.0000	0.9992***
	(0.3555)	(0.0848)	(0.0003)	(0.0000)	(0.0003)
ISHMT27	0.6937***	0.3972***	1.0017***	1.0000	0.9993
	(0.2046)	(0.0518)	(0.0005)	(0.0000)	(0.0005)
ISHMT28	0.0421	0.4383***	1.0001	1.0001**	1.0009**
	(0.1502)	(0.0498)	(0.0004)	(0.0000)	(0.0004)
ISHMT29	0.4153***	0.3435***	1.0021***	0.9999	0.9996

Table continues in following page.

Table B.1: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
	(0.0751)	(0.0323)	(0.0004)	(0.0001)	(0.0004)
ISHMT30	0.1937**	0.7001***	0.9999*	1.0004***	0.9998**
	(0.0872)	(0.1259)	(0.0000)	(0.0001)	(0.0001)
ISHMT31	0.6849***	1.0468***	1.0001	1.0004***	1.0001
	(0.0699)	(0.0674)	(0.0001)	(0.0001)	(0.0001)
ISHMT32	0.4664***	0.3013***	0.9997*	0.9992***	1.0001*
	(0.0355)	(0.0400)	(0.0002)	(0.0002)	(0.0001)
ISHMT33	0.1249***	0.1348***	1.0000	1.0000	1.0000
	(0.0431)	(0.0439)	(0.0000)	(0.0002)	(0.0000)
ISHMT34	0.4034***	0.3014***	1.0012***	0.9980***	0.9997***
	(0.0174)	(0.0169)	(0.0003)	(0.0005)	(0.0001)
ISHMT35	0.1969***	0.0385***	1.0000	0.9943***	1.0000
	(0.0137)	(0.0142)	(0.0002)	(0.0007)	(0.0002)
ISHMT36	0.2319***	0.2978***	1.0004**	1.0014***	1.0001*
	(0.0168)	(0.0169)	(0.0002)	(0.0005)	(0.0001)
ISHMT37	0.3648***	0.2723***	0.9999	0.9990***	1.0000
	(0.0235)	(0.0247)	(0.0002)	(0.0004)	(0.0001)
ISHMT38	0.1266***	0.0833	0.9998**	0.9999	1.0001
	(0.0433)	(0.0612)	(0.0001)	(0.0002)	(0.0001)
ISHMT39	0.1555***	0.1264***	0.9992***	0.9997	1.0001
	(0.0232)	(0.0314)	(0.0001)	(0.0005)	(0.0002)
ISHMT40	0.4595***	0.6002***	1.0000	1.0015***	1.0000
	(0.0243)	(0.0254)	(0.0003)	(0.0004)	(0.0001)
ISHMT41	0.7076***	0.5908***	0.9989**	0.9980***	1.0002**
	(0.0193)	(0.0207)	(0.0005)	(0.0005)	(0.0001)
ISHMT42	0.4153***	0.4218***	1.0005***	1.0000	1.0000
	(0.0435)	(0.0379)	(0.0001)	(0.0002)	(0.0001)
ISHMT43	0.4479***	0.4303***	1.0001	0.9997	1.0000
	(0.0188)	(0.0193)	(0.0003)	(0.0005)	(0.0000)
ISHMT44	0.2955***	0.2503***	1.0001	0.9996	1.0000
	(0.0266)	(0.0270)	(0.0001)	(0.0003)	(0.0000)
ISHMT45	0.7686***	0.8144***	0.9999	1.0006	1.0000
	(0.0213)	(0.0221)	(0.0005)	(0.0004)	(0.0000)
ISHMT46	0.6415***	0.6906***	1.0004***	1.0001	1.0000
	(0.0704)	(0.0592)	(0.0001)	(0.0001)	(0.0001)
ISHMT47	0.5503***	0.3778***	0.9995***	0.9991***	1.0002**
	(0.0333)	(0.0384)	(0.0002)	(0.0003)	(0.0001)
ISHMT48	0.3510***	0.2873***	1.0011***	0.9983***	0.9998**
	(0.0152)	(0.0150)	(0.0003)	(0.0006)	(0.0001)

Table continues in following page.

Table B.1: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
ISHMT49	0.2670***	0.1772***	0.9999	0.9984***	1.0000
	(0.0194)	(0.0203)	(0.0002)	(0.0005)	(0.0001)
ISHMT50	0.3487***	0.2641***	1.0041***	0.9984***	0.9990***
	(0.0182)	(0.0157)	(0.0004)	(0.0005)	(0.0003)
ISHMT51	0.2823***	0.2624***	1.0009***	0.9997	0.9999
	(0.0199)	(0.0190)	(0.0002)	(0.0004)	(0.0001)
ISHMT52	0.4371***	0.3131***	0.9997**	0.9995**	1.0001
	(0.0395)	(0.0454)	(0.0001)	(0.0002)	(0.0001)
ISHMT53	0.4794***	0.4423***	0.9998	0.9997	1.0000
	(0.0249)	(0.0264)	(0.0002)	(0.0003)	(0.0000)
ISHMT54	0.2594***	0.1714***	1.0005***	0.9990***	0.9998**
	(0.0231)	(0.0225)	(0.0002)	(0.0004)	(0.0001)
ISHMT55	0.2198***	0.2562***	0.9998	1.0003	1.0000
	(0.0272)	(0.0295)	(0.0001)	(0.0003)	(0.0000)
ISHMT56	0.4056***	0.3142***	1.0007***	0.9991***	0.9998**
	(0.0249)	(0.0239)	(0.0002)	(0.0003)	(0.0001)
ISHMT57	0.0026	-0.0841***	1.0000	0.9977***	0.9999
	(0.0156)	(0.0160)	(0.0000)	(0.0006)	(0.0001)
ISHMT58	0.1798***	-0.1437***	1.0002**	0.9988***	0.9997***
	(0.0385)	(0.0357)	(0.0001)	(0.0002)	(0.0001)
ISHMT59	-0.0809***	-0.2216***	0.9997***	0.9974***	0.9995***
	(0.0178)	(0.0170)	(0.0001)	(0.0005)	(0.0001)
ISHMT60	0.1182***	0.0041	0.9999*	0.9994**	1.0001
	(0.0343)	(0.0387)	(0.0001)	(0.0003)	(0.0001)
ISHMT61	-0.1122***	-0.2568***	0.9993***	0.9973***	0.9990***
	(0.0178)	(0.0161)	(0.0001)	(0.0005)	(0.0002)
ISHMT62	0.9671***	1.0283***	1.0005	1.0003	1.0000
	(0.0328)	(0.0326)	(0.0004)	(0.0002)	(0.0000)
ISHMT63	0.5658***	0.5955***	0.9986***	1.0003	0.9999
	(0.0227)	(0.0265)	(0.0003)	(0.0004)	(0.0001)
ISHMT64	0.2986***	0.2153***	1.0003*	0.9990***	0.9999
	(0.0223)	(0.0224)	(0.0002)	(0.0004)	(0.0001)
ISHMT65	0.3375***	0.3503***	1.0008***	1.0001	1.0000
	(0.0294)	(0.0267)	(0.0002)	(0.0003)	(0.0001)
ISHMT66	0.1060***	-0.0492	0.9988***	0.9975***	1.0018***
	(0.0194)	(0.0377)	(0.0002)	(0.0007)	(0.0005)
ISHMT67	0.5208***	0.3471***	1.0005***	0.9995***	0.9998**
	(0.0443)	(0.0401)	(0.0002)	(0.0002)	(0.0001)
ISHMT68	0.0591**	-0.0655***	1.0002**	0.9987***	0.9996***

Table continues in following page.

Table B.1: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
	(0.0236)	(0.0217)	(0.0001)	(0.0003)	(0.0001)
ISHMT69	0.1455***	0.0739***	0.9995***	0.9988**	1.0003**
	(0.0187)	(0.0218)	(0.0001)	(0.0005)	(0.0001)
ISHMT70	0.1353***	0.0370**	1.0009***	0.9986***	0.9994***
	(0.0203)	(0.0176)	(0.0002)	(0.0004)	(0.0002)
ISHMT71	0.4876***	0.3153***	1.0000	0.9997*	1.0000
	(0.0618)	(0.0631)	(0.0001)	(0.0001)	(0.0000)
ISHMT72	0.4052***	0.1034**	1.0003**	0.9995***	0.9998**
	(0.0577)	(0.0511)	(0.0001)	(0.0001)	(0.0001)
ISHMT73	0.7561***	0.6918***	1.0011***	0.9993*	0.9999
	(0.0234)	(0.0230)	(0.0004)	(0.0004)	(0.0001)
ISHMT74	0.3847***	0.2964***	1.0003**	0.9996*	0.9999
	(0.0380)	(0.0364)	(0.0001)	(0.0002)	(0.0001)
ISHMT75	0.4681***	0.3526***	1.0006***	0.9997*	0.9999*
	(0.0440)	(0.0387)	(0.0002)	(0.0002)	(0.0001)
ISHMT76	0.2162***	0.0182	1.0012***	0.9978***	0.9989***
	(0.0235)	(0.0198)	(0.0002)	(0.0003)	(0.0002)
ISHMT77	0.0856***	0.0327	0.9999**	0.9997	1.0001
	(0.0300)	(0.0358)	(0.0001)	(0.0003)	(0.0001)
ISHMT78	0.3533***	0.1826***	0.9999	0.9996**	1.0001
	(0.0486)	(0.0542)	(0.0001)	(0.0002)	(0.0000)
ISHMT79	03341***	0.2062***	0.9999	0.9964***	1.0000
	(0.0150)	(0.0156)	(0.0003)	(0.0006)	(0.0001)
ISHMT80	1.4094***	1.1943***	1.0005	0.9984***	0.9999
	(0.0278)	(0.0272)	(0.0007)	(0.0003)	(0.0001)
ISHMT81	1.1299***	1.1659***	0.9964***	1.0005	0.9999
	(0.0210)	(0.0248)	(0.0007)	(0.0004)	(0.0001)
ISHMT82	0.7126***	0.5867***	0.9986***	0.9987***	1.0002**
	(0.0242)	(0.0273)	(0.0004)	(0.0004)	(0.0001)
ISHMT83	0.6003***	0.4962***	1.0007	0.9971***	0.9999
	(0.0150)	(0.0154)	(0.0005)	(0.0006)	(0.0001)
ISHMT84	0.6150***	0.4738***	1.0003**	0.9999	0.9999
	(0.0733)	(0.0648)	(0.0001)	(0.0001)	(0.0000)
ISHMT85	0.5586***	0.2782***	1.0011***	0.9984***	0.9995***
	(0.0317)	(0.0287)	(0.0002)	(0.0003)	(0.0001)
ISHMT86	0.6819***	0.2799***	1.0003	0.9977***	0.9998
	(0.0315)	(0.0318)	(0.0003)	(0.0003)	(0.0002)
ISHMT87	0.1998***	0.0755***	1.0001	0.9982***	0.9999
	(0.0202)	(0.0208)	(0.0001)	(0.0004)	(0.0001)

Table continues in following page.

Table B.1: (continued)

			Decomposition (as ratio)		
	2007/08	2014/15	Characteristics	Coefficients	Interaction
ISHMT88	0.3645***	0.2872***	1.0000	0.9981***	1.0000
	(0.0157)	(0.0163)	(0.0003)	(0.006)	(0.0001)
ISHMT89	0.6281***	0.3417***	1.0013***	0.9980***	0.9994***
	(0.0289)	(0.0265)	(0.0003)	(0.0003)	(0.0002)
ISHMT90	0.3525***	0.4006***	1.0009***	1.0002	1.0001
	(0.0333)	(0.0288)	(0.0002)	(0.0002)	(0.0001)
ISHMT91	0.1251***	0.1354***	1.0002***	1.0001	1.0000
	(0.0316)	(0.0277)	(0.0001)	(0.0003)	(0.0001)
ISHMT92	0.4529***	0.4154***	1.0000	0.9997	1.0000
	(0.0286)	(0.0300)	(0.0002)	(0.0003)	(0.0000)
ISHMT93	0.2285***	0.2116***	1.0004*	0.9995	1.0000
	(0.0142)	(0.0146)	(0.0002)	(0.0006)	(0.0000)
ISHMT94	0.3004***	0.4168***	0.9996***	1.0006**	0.9999*
	(0.0341)	(0.0411)	(0.0001)	(0.0003)	(0.0001)
ISHMT95	0.1678***	0.2007***	0.9999	1.0004	1.0000
	(0.0232)	(0.0248)	(0.0001)	(0.0004)	(0.0000)
ISHMT96	0.5659***	0.4150***	0.9995***	0.9995**	1.0001*
	(0.0427)	(0.0526)	(0.0002)	(0.0002)	(0.0001)
ISHMT97	0.4394***	0.5390***	0.9999	1.0003	1.0000
	(0.0435)	(0.0474)	(0.0001)	(0.0002)	(0.0000)
ISHMT98	0.2602***	0.2687***	0.9987***	1.0001	1.0000
	(0.0213)	(0.0279)	(0.0002)	(0.0005)	(0.0002)
ISHMT99	0.5013***	0.5178***	0.9971***	1.0003	0.9999
	(0.0172)	(0.0207)	(0.0004)	(0.0006)	(0.0002)
ISHMT100	-0.0991***	-0.3130***	1.0004***	0.9979***	1.0008***
	(0.0248)	(0.0326)	(0.0001)	(0.0004)	(0.0002)
ISHMT101	0.3494***	0.4439***	0.9992***	1.0010**	0.9998**
	(0.0244)	(0.0286)	(0.0002)	(0.0004)	(0.0001)
ISHMT102	0.2388***	0.2444***	0.9994***	1.0001	1.0000
	(0.0233)	(0.0269)	(0.0001)	(0.0004)	(0.0001)
ISHMT103	0.6276***	0.7591***	0.9998*	1.0001	1.0000
	(0.0781)	(0.0978)	(0.0001)	(0.0001)	(0.0000)
ISHMT104	0.5272**	0.7554**	1.0000	1.0000	1.0000
	(0.2147)	(0.3379)	(0.0000)	(0.0000)	(0.0000)
ISHMT105	0.7845	1.0752**	1.0000	1.0000	1.0000
	(0.6069)	(0.4777)	(0.0000)	(0.0000)	(0.0000)
ISHMT106	0.4656***	0.4309***	1.0000	1.0000	1.0000
	(0.1298)	(0.1303)	(0.0000)	(0.0001)	(0.0000)
ISHMT107	0.2399*	0.4130***	1.0001	1.0001	1.0000

Table continues in following page.

Table B.1: (continued)

			Decomposition (as ratio)		
	2007/08	2014/15	Characteristics	Coefficients	Interaction
	(0.1242)	(0.1011)	(0.0000)	(0.0001)	(0.0000)
ISHMT108	0.7920***	0.8287***	1.0000	1.0000	1.0000
	(0.1156)	(0.1150)	(0.0001)	(0.0001)	(0.0000)
ISHMT109	0.5508***	0.2835***	1.0006***	0.9990***	0.9997***
	(0.0394)	(0.0366)	(0.0002)	(0.0002)	(0.0001)
ISHMT110	0.7767***	0.6737***	0.9998	0.9989***	1.0000
	(0.0241)	(0.0256)	(0.0004)	(0.0004)	(0.0001)
ISHMT111	0.0561***	-0.0727***	0.9998***	0.9955***	1.0005***
	(0.0139)	(0.0152)	(0.0001)	(0.0007)	(0.0001)
ISHMT112	0.1732***	0.1397***	1.0001	0.9988*	1.0000
	(0.0137)	(0.0141)	(0.0002)	(00007x)	(0.0000)
ISHMT113	-0.5909**	-0.3513	1.0000	1.0000	1.0000
	(0.2718)	(0.4787)	(0.0000)	(0.0000)	(0.0000)
ISHMT114	0.1311***	0.0647***	1.0010***	0.9925***	0.9995***
	(0.0085)	(0.0085)	(0.0002)	(0.0013)	(0.0001)
ISHMT115	0.5084***	0.5502***	1.0003**	1.0001	1.0000
	(0.0537)	(0.0484)	(0.0001)	(0.0001)	(0.0000)
ISHMT116	0.1957***	0.1506***	0.9998	0.9991*	1.0000
	(0.0177)	(0.0187)	(0.0001)	(0.0005)	(0.0000)
ISHMT117	0.7646***	0.9425***	0.9996	1.0013***	0.9999
	(0.0279)	(0.0302)	(0.0003)	(0.0003)	(0.0001)
ISHMT118	1.1389***	1.0693***	1.0001	0.9993*	1.0000
	(0.0249)	(0.0257)	(0.0006)	(0.0004)	(0.0000)
ISHMT119	1.2447***	1.0873***	0.9992	0.9990***	1.0001
	(0.0302)	(0.0331)	(0.0005)	(0.0003)	(0.0001)
ISHMT120	0.5973***	0.5402***	1.0010*	0.9981***	0.9999
	(0.0142)	(0.0143)	(0.0006)	(0.0007)	(0.0001)
ISHMT121	0.8667***	0.7511***	1.0002	0.9999	1.0000
	(0.0705)	(0.0660)	(0.0002)	(0.0001)	(0.0000)
ISHMT122	-0.1403***	0.0378*	0.9999	1.0025***	1.0001
	(0.0212)	(0.0215)	(0.0001)	(0.0004)	(0.0001)
ISHMT123	0.6030***	0.6182***	1.0015***	1.0003	1.0000
	(0.0186)	(0.0181)	(0.0004)	(0.0005)	(0.0001)
ISHMT124	1.3337**	(omitted)	1.0000	1.0000	1.0000
	(0.0069)		(0.0000)	(0.0000)	(0.0000)
ISHMT125	0.1600***	0.1536***	0.9999	1.0000	1.0000
	(0.0368)	(0.0409)	(0.0001)	(0.0002)	(0.0000)
ISHMT126	0.0790**	-0.2029***	1.0001	0.9985***	0.9998**
	(0.0326)	(0.0315)	(0.0000)	(0.0003)	(0.0001)

Table continues in following page.

Table B.1: (continued)

	2007/08	2014/15	Decomposition (as ratio) Characteristics	Coefficients	Interaction
ISHMT127	0.0884**	0.2066***	0.9998**	1.0006*	0.9997*
	(0.0349)	(0.0538)	(0.0001)	(0.0003)	(0.0002)
ISHMT128			(omitted)		
ISHMT129	0.3872	-2.8191***	1.0000	0.9999	1.0001
	(0.3508)	(0.6755)	(0.0000)	(0.0001)	(0.0001)
ISHMT130	0.1965***	0.1311***	0.9992***	0.9969***	1.0003***
	(0.0115)	(0.0125)	(0.0002)	(0.0008)	(0.0001)
First Diagnosis - total effect			1.0161***	0.9087***	0.9912***
			(0.0027)	(0.0080)	(0.0019)
Comorbidities	0.0734***	0.0811***	1.0517***	1.0093***	1.0053***
	(0.0022)	(0.0020)	(0.0017)	(0.0035)	(0.0020)
Type of Provider					
Large Provider	Reference Category		1.0016***	0.9963	1.0005
			(0.0005)	(0.0043)	(0.0006)
Medium Provider	-0.0050	0.0098	1.0008***	1.0013	0.9999
	(0.0060)	(0.0065)	(0.0002)	(0.0034)	(0.0003)
Small Provider	-0.0042	-0.0147	1.0004***	0.9977	1.0002
	(0.0081)	(0.0090)	(0.0001)	(0.0015)	(0.0001)
Specialist Provider	0.1803***	0.1666***	1.0000	0.9993	1.0000
	(0.0150)	(0.0158)	(0.0001)	(0.0006)	(0.0000)
Teaching Provider	0.0719***	0.0731***	1.0022***	0.9980	0.9995
	(0.0064)	(0.0064)	(0.0006)	(0.0027)	(0.0007)
Other Provider	-0.0305	0.0390**	0.9988	1.0001	1.0011
	(0.0575)	(0.0191)	(0.0009)	(0.0001)	(0.0009)
Type of Provider - total effect			1.0038***	0.9928	1.0011
			(0.0011)	(0.0104)	(0.0012)
Constant	6.5440***	6.6594***	-	1.0523***	-
	(0.0126)	(0.0133)		(0.0136)	
Adjusted R-squared	0.6630	0.6666			
N	66,079	75,343		141,422	

Standard errors in parenthesis. ***, **, * represent 1%, 5% and 10% significance, respectively.

B.2. Counterfactual Decomposition

Table B.2: Counterfactual Decomposition of Differences in Distributions - Log(Hospital Costs) between 2007/08 and 2014/15 - 130 diagnosis groups

Quantile	Differences between the observable distributions	Effects of characteristics	Effects of coefficients
0.1	-0.109	-0.018	-0.091
	[-0.188, -0.100]	[-0.026, -0.011]	[-0.100, -0.082]
0.2	-0.032	-0.014	-0.018
	[-0.043, -0.021]	[-0.021, -0.007]	[-0.027, -0.009]
0.3	0.019	0.002	0.017
	[0.007, 0.030]	[-0.005, 0.008]	[0.009, 0.025]
0.4	0.050	0.008	0.042
	[0.037, 0.063]	[0.001, 0.015]	[0.033, 0.051]
0.5	0.081	0.016	0.065
	[0.067, 0.095]	[0.009, 0.023]	[0.055, 0.074]
0.6	0.096	0.027	0.069
	[0.082, 0.111]	[0.019, 0.035]	[0.060, 0.078]
0.7	0.110	0.044	0.066
	[0.094, 0.126]	[0.033, 0.054]	[0.056, 0.076]
0.8	0.138	0.081	0.057
	[0.118, 0.158]	[0.068, 0.094]	[0.043, 0.072]
0.9	0.216	0.184	0.032
	[0.195, 0.238]	[0.162, 0.207]	[0.012, 0.051]
N 2007/08		66,079	
N 2014/15		75,343	

Pointwise Confidence Interval in parenthesis.

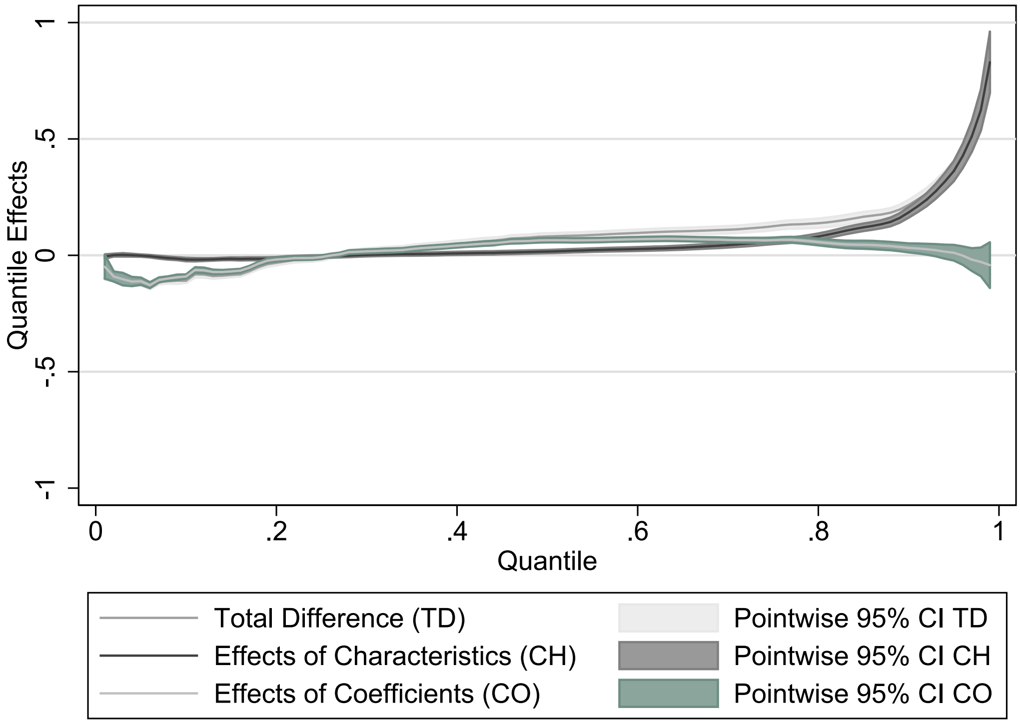


Figure B.1: Counterfactual Decomposition of Differences in Distributions - Log(Hospital Costs) between 2007/08 and 2014/15 - 130 diagnosis groups