MONASH University

MONASH BUSINESS SCHOOL

ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

A Panel Data Analysis of Hospital Variations in Length of Stay for Hip Replacements: Private versus Public

Yan Meng, Xueyan Zhao, Xibin Zhang and Jiti Gao

December 2017

Working Paper 20/17

AACSB ACCREDITED EFMD EQUIS ACCREDITED ASSOCIATION OF MBAS AMBA ACCREDITED

A Panel Data Analysis of Hospital Variations in Length of Stay for Hip Replacements: Private versus Public

Yan Meng[1], Xueyan Zhao, Xibin Zhang and Jiti Gao[2]

Department of Econometrics and Business Statistics, Monash University, Australia

(Working Paper Only)

November 27, 2017

ABSTRACT: Inequality between private and public patients in Australia has been an ongoing concern due to its two tiered insurance system. This paper investigates the variations in hospital length of stay for hip replacements using Victorian Admitted Episodes Dataset from 2003/2004 to 2014/2015, employing a Bayesian hierarchical random coefficient model with trend allowing for structural break. We find systematic differences in the length of stay between public and private hospitals, after observable patient complexity is controlled. This suggests shorter stay in public hospitals due to pressure from Activity-based funding scheme, and longer stay in private system due to potential moral hazard. Our counterfactual analysis shows that public patients stay 1.4 days shorter than private in 2014, which leads to the 'quicker but sicker' concern that is commonly voiced by the public. We also identify widespread variations among individual hospitals. Sources for such variation warrant closer investigation by policy makers.

KEYWORDS: Gibbs sampler, hierarchical random coefficients, length of stay, hospital ranking

JEL Classification: C11, H51, I14

[1]Address: Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia. Telephone: +61 3 99034535. Fax: +61 3 99032007. Email: *yan.meng@monash.edu*

[2]This author's research was supported by the Australian Research Council Discovery Grants Program under Grant Numbers: DP150101012 and DP170104421.

1 Introduction

As a key component of Australian comprehensive system of health services, hospitals make a great contribution to the national health outcomes. Inequality among patients has become a major concern for the government and commentators in terms of affordability, access to services, waiting time and the quality of care. Patients, if treated in public hospitals, are covers by the universal, tax-financed health insurance scheme, Medicare, which provides rebates against the cost of medical fees. Private health insurance covers patients in private hospitals or else the patients need to pay for their treatment out-of-pocket. Patients can also choose to be treated in public hospitals as private patients, in which case Medicare will pay partial amount of the cost for the stay with the rest being covered by the private health insurance or out-of-pocket payments. It is clear that different types of hospitals are driven by different operational motives, which inevitably lead to a common concern that public patients receive less quality of care relative to private patients. It has been found that Victorian private hospitals persistently outperform public hospitals (Jensen et al., 2009; Palangkaraya and Yong, 2013). In particular, it is believed that public patients are allowed to stay for a much shorter time in hospitals due to the overwhelming demand for hospital beds, long waiting time and funding pressure from the government. Concurrently there is also a concern for any 'moral hazard' in private hospitals, given that insurance will pay for the cost of the stay.

The health care system in Australia is comprehensive and largely publicly funded. When the patient LOS lies within the recommended lower and upper boundaries (see Table A.1) in the Weighted Inlier Equivalent Separation (WIES) table, the current activity-based funding (ABF) scheme will essentially pay a fixed 'price' per episode to public hospitals, with the price being determined by Australian Refined Diagnosis Related Groups (AR-DRG). Due to the financial incentives in this funding regime, public hospitals have little incentive to keep the patients longer in order to enjoy the largest price-cost margin (Hall, 2010). Theoretically speaking, there should be a length of stay (LOS) that is required for given surgery and given patient complexity regardless of whether the payment is public or private.

There is also more interest recently in giving more information and greater choice to patients for choosing hospitals. For this reason, benchmarking and comparison across individual hospitals after controlling for patient complexity is also important. Patients with more complex conditions normally stay in hospitals longer and therefore consume more resources. However a shorter LOS is favored by the hospitals because it reduces the cost and more importantly frees up beds quickly and enables more patients to be treated given limited number of beds available in each hospital.

Among many surgeries performed in hospitals each year, total hip replacement (THR) has been described as "the operation of the century" (Learmonth et al., 2007). As one of the most common reasons for the hip joint damage that causes pain, the hospitalization rate for patients with osteoarthritis increased from 362 per 100,000 population in 2005 to 415 per 100,000 population in 2014 (AIHW, 2012). The rate of hip replacement, as one of the most frequently performed surgeries in Australia, rose by 22 percent over the same period (AIHW, 2012). Hospital expenditure is the largest single contributor to the growth in national public spending, among which expenditure on hip replacements outstrips most other surgical procedures in the health sector in Australia due to high volumes, as there has always been a long waiting time for this surgery.

The average LOS (ALOS) of a hospital stay for hip replacement has decreased significantly in recent years but the discharge of patients to non-home settings and the rate of readmission for complications or referrals to skilled care facilities has increased (Cram et al., 2011). Stays that are overly short may lead to a knock-on adverse effect on the quality of care experienced by the patients and therefore cause diminishing patient outcomes (Martin et al., 2016). Moreover, patients discharged prematurely are placed at a greater risk of subsequent readmission to hospitals (Bueno et al., 2010; Carey, 2015).

With the effort for better management of hospital costs and improvement of healthcare quality, organizations have a growing interest in the determinants and variations of hospital LOS. To study this, different approaches have been used in the literature including linear regression using log-transformed LOS (Draper and Luscombe, 1998; Eastwood et al., 1999; Rosen et al., 1999), quantile regression for several estimates of differences in LOS (Hauck and Hollingsworth, 2008), generalized linear models (GLM) with a log link and Gaussian, Poisson, negative binomial or gamma distributions to characterize the relationship between the conditional mean and variance (Austin et al., 2002; Castelli et al., 2015) and two-component mixture models to accommodate both inherent correlation and heterogeneity (Thompson et al., 1998; Yau et al., 2003). More recent and extensive overviews can be found in Jones (2000) and Moran and Solomon (2012).

A substantial empirical literature has emerged that examines variations in hospital LOS. Studies have found that LOS is associated with patient level control variables such as age, gender, marital status, socioeconomic index (Cheng et al., 2014; Correia and Waitzberg, 2003; Husted et al., 2008; Segal et al., 2009), source of referral (Liu et al., 2001) and discharge destination (Husted et al., 2010). Apart from these, hospital level characteristics also play a crucial role such as teaching status, size and location of the hospitals (Hauck and Hollingsworth, 2008). Several proxies were used to measure the level of severity, including the patient's medical history (Wright et al., 2003), admission to ICU (Morales et al., 2003) and the overall Charlson Comorbidity Index (Charlson et al., 1994; Librero et al., 1999; Sundararajan et al., 2004). However, a large proportion of variation in LOS is sometimes not captured by the performance data and research on the unobserved factors is scarce.

Recent literature in health economics has also recognized the importance of data structure in the analysis (see Adams et al., 2003; Castelli et al., 2013; Koop et al., 1997; Olsen and Street, 2008; Rosenberg et al., 1999; Sørensen et al., 2009; Zhang et al., 2013). Variations in the individual and group characteristics can be studied simultaneously. The dependence between the parameters are defined by the hierarchical structure in a way that enough parameters are included in the model to fit the data and in the meantime overfitting is avoided. The standard approach uses mixed effect models to describe the system with multiple components of variation, in which the fixed effects represent quantities directly, whereas the random effects are quantities sampled from a population about which inference is desired (Schmid and Brown, 2000).

The objective of this paper is to study the effects of observed and unobserved patient and hospital level factors on the LOS for hip surgeries. Due to the hierarchical structure of the data, we employ the Bayesian approach by specifying a three-level random coefficient model. We use a rich administrative hospital dataset from the state of Victoria in Australia for the period of 2003 to 2014, and we pay particular attention to modeling unobserved hospital level factors. We explore the multi-level panel structure of the data to estimate the unobserved heterogenous hospital effects both in the intercept and in the slope for patient complexity. The comparative information is an important input for policy makers to benchmark hospitals, in order to improve the efficiency and quality of patient care in all hospitals. We also include a trend function with a structural break allowing for changes in medical technologies for hip surgeries over time, and a structural change due to faulty hip material.

A particular focus of this paper is to identify any systematic differences between the public and private hospitals to shed light on inequality and potential moral hazard issues mentioned above. We conduct a counterfactual analysis by predicting the LOS for the same ”average” patient in all hospitals, in particular in private versus public hospitals. The use of modern Bayesian approach also allows us to present uncertainties in the results in the form of posterior densities for the measures of interest, including hospital effects.

The paper proceeds as follows. Section 2 describes the data and the construction of the variables. Section 3 proposes a Bayesian method to estimate the random intercept hierarchy model. Estimation results and findings are given in Section 4 and Section 5 concludes the paper. Appendices A and B include some summary statistics, Appendix C summaries the Gibbs sampling procedure, and Appendix D gives some additional table and figures.

2 Data

We use the Victorian Admitted Episodes Dataset (VAED) for patients who had a hip replacement over the period from 2003/2004 to 2014/2015. The VAED comprises demographic, clinical and administrative details for every admitted episode of care occurring in all public and private hospitals, rehabilitation centers, extended care facilities and day procedure centers in Victoria. It is commonly used to administer the casemix funding system for public hospitals in Victoria, to assess the utilization and performance of health services and to monitor population morbidity in order to inform health policy for possible improvements of health services.

Table 1: Variable definitions and summary statistics

Variable	Definition	Percentage	Mean
y	log of length of stay for hip replacement episode, measured in log(days)		1.91872
age	age of the patients		70.9727
male	1 if male, 0 otherwise	40.26	0.4026
married	1 if married or de facto, 0 otherwise	59.74	0.5974
int_care	1 if the patient has been to ICU and/or CCU during the episode, 0 otherwise	7.98	0.0798
emergency	1 if emergency patient, 0 for elective patient	22.02	0.2202
pub_pub	1 if the patient is public and is admitted to a public hospital, 0 otherwise	38.89	0.3889
pri_pub	1 if the patient is private and is admitted to a public hospital, 0 otherwise	6.82	0.0682
pri_pri	1 if the patient is private and is admitted to a private hospital, 0 otherwise	54.29	0.5429
severe	1 if DRG code is I03A (major complexity), 0 if the code is I03B (minor complexity)	15.90	0.1590
major	1 if the hospital had more than 1000 hip replacement episodes in sample period, 0 otherwise	33.28	0.3328
metro	1 if the hospital is in metropolitan area, 0 for rural	74.61	0.7461
sep_T	1 if separation and transfer to acute/extended care/rehab/geriatric center, 0 otherwise	33.87	0.3387
sep_N	1 if separation and tansfer to aged care residential facility, 0 otherwise	2.31	0.0231
sep_S	1 if change in care type within this hospital at separation, 0 otherwise	11.73	0.1173
sep_H	1 if separation to private residence/accomodation, 0 otherwise	52.09	0.5209
t	year in which the patient was admitted		8.0728

Table 2: Summary statistics for LOS (in days) in each group

Group	Number of observations	Summary statistics for LOS (in days) in each group				
		average LOS	median LOS	SD	Min	Max
LOS	59,915	8.1	7	6.7	1	174
age						
<40	969	6.2	5	1.9	2	7
40-50	2,930	6.1	5	5.4	1	68
50-60	8,222	6.4	6	5.6	1	88
60-70	15,217	7.0	6	4.7	1	136
70-80	17,451	8.4	7	5.7	1	174
80-90	12,872	10.3	8	7.8	1	170
>90	2,254	11.5	9	8.3	2	80
t_{2003}	2,340	12.6	10	10.9	1	174
t_{2014}	8,323	6.0	5	3.8	1	114
male	24,121	7.8	6	6.5	1	174
female	35,794	8.4	7	6.8	1	170
married	35,792	7.7	6	6.0	1	174
unmarried	24,123	8.9	7	7.5	1	170
int_care	4,779	12.1	9	10.7	1	174
no int_care	55,136	7.8	6	6.1	1	170
emergency	13,191	11.8	9	9.9	1	174
elective	46,724	7.1	6	5.0	1	141
pub_pub	23,302	8.4	6	7.4	1	174
pri_pub	4,084	9.8	8	8.9	1	170
pri_pri	32,529	7.8	7	5.8	1	147
severe	9,524	11.7	9	10.3	1	174
non-severe	50,391	7.5	6	5.5	1	167
major	48,975	8.2	7	6.9	1	174
non-major	10,940	8.0	7	5.6	1	116
metro	44,702	8.4	7	7.0	1	174
non-metro	15,213	7.3	6	5.7	1	141
sep_T	20,293	9.6	8	7.3	1	174
sep_N	1,383	11.8	8	12.2	1	170
sep_S	7,030	7.6	6	6.7	1	122
sep_H	31,209	7.2	6	5.6	1	141

Our sample consists of 59,915 episodes from 20 public and 20 private hospitals over 12 years. We exclude those whose admission type was maternity or statistical (change in care type within the same hospital at admission), so that the patients were admitted either from the waiting list or through the emergency department. We drop sameday stays as hospital preparations and patient preferences for such stays are very different from overnight stays. We further exclude episodes from public hospitals that are not wies-fundable and episodes in which the patients' region of residence are interstate or missing. Patient type comes in five categories in the dataset - private, public, compensable, Department of Veterans' Affairs (DVA) and ineligible. Since private patients are defined to be those who are charged or for whom a charge is raised for a third party payer, we combined DVA and compensable patient into private and dropped the ineligible ones. Among the remaining hospitals, those that performed less than 400 hip replacements in the sample period were excluded for sample robustness. To make sure enough information was provided to individual hospitals in each year, two hospitals with missing values were removed. A cross tabulation for hospitals and years is given in Table B.1. We ended up with 20 public hospitals and 20 private hospitals. Table 1 shows the definition of all variables used.

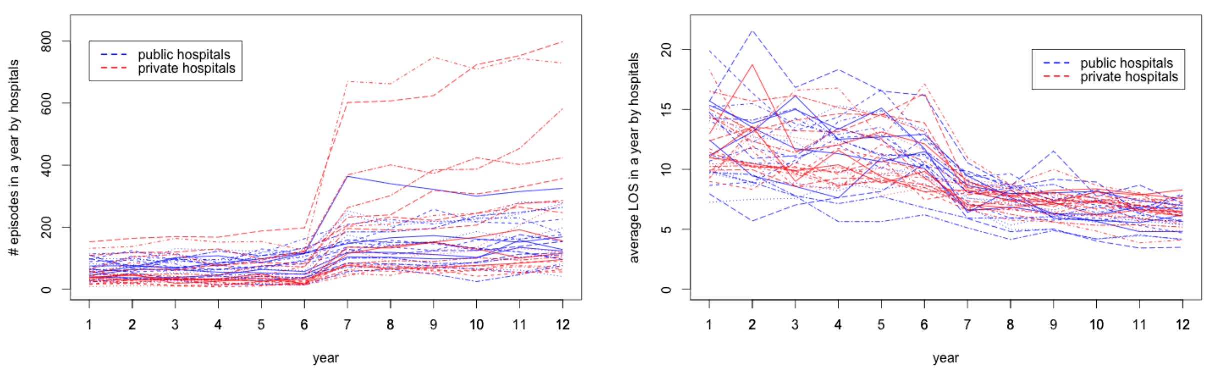
The dependent variable is the log of LOS of hip replacement patients in the hospitals. We choose hip replacement episodes based on the AR-DRG code. AR-DRG is a patient classification system which provides a clinically meaningful way of relating the number and type of patients treated in hospital to the resources required by the hospital. The classification categorizes acute admitted patient episodes of care into groups with similar conditions and similar usage of hospital resources, using information in the hospital morbidity record such as diagnoses, procedures and demographic characteristics of the patients.[3] Hence, we introduced the dummy variable 'severe' to differentiate the complexity of the surgery.

Several variables on patient demographic characteristics are selected to control for differences in hospital care need, including gender, age and marital status. If a patient is transferred to another hospital, it is recorded as a new episode. Since we do not know at what stage during the admission that the patient is transferred and for what the reason the patient is transferred, we cannot link several episodes, especially when the patient or hospital type changes. We try to control this by taking account the information of

[3]The AR-DRG classification is partly hierarchical, with 23 Major Diagnostic Categories (MDCs) into which the 661 AR-DRGs can be grouped. I03A and I03B are for hip replacement with major and minor complexity respectively, where 'I' is one of the MDCs standing for diseases and disorders of the musculoskeletal system and connective tissue, 'A' indicates the highest consumption of resources and 'B' for second highest.

the patient's discharge destination. To control for hospital characteristics, we included 'metro' and 'major' in the regression. The key explanatory variables of interest are the patient and hospital type variables based on the account payment type in each episode. We therefore categorize patient and hospital types into public patients in public hospitals, private patients in public hospital and private patients in private hospitals. Table 2 gives the reader a brief idea of LOS (in days) in each category.

We plot the number of episodes and ALOS in each hospital across all years, shown in Figure 1a and Figure 1b respectively. Looking through the sample period, it is obvious that the total number of episodes jumped up dramatically from the seventh year in most hospitals. The number of episodes in a few private hospitals doubled or even tripled the amount in previous years, indicated by the dotted red lines. Accordingly, ALOS has been decreasing relatively quicker from the beginning of the sample period till 2009 and after that there was not much room for individual hospitals to shorten the LOS further.



(a) Episodes in each hospital across time (b) ALOS in each hospital across time

Figure 1: Number of episodes and average LOS in each hospital, 2003-2014

3 A panel data model with random coefficients

3.1 Model

Hierarchical models analyze variability arising at distinct levels within complex data. Analysis of patients in hospitals is a natural application of hierarchical modeling in the health field. For example, hospital characteristics such as quality of services or managerial performance may impose distinct effects on the cost of treating patients. For this reason, we specify a hierarchical random coefficient model to relate the hospital length of stay during a hip replacement episode to observable patient and hospital characteristics, random error components of hospital effects, time effects and the remaining random variations.

Let y_{iht} denote the logarithm of length of stay for the i-th episode in the h-th hospital and the t-th year. Our model can be specified as the following:

$$\begin{aligned}
&y_{iht} = X_{iht}\beta_h + Z_{iht}^\top\gamma + \alpha_h + \delta_1 t + (\delta_2 - \delta_1)tI(t \geq 7) + e_{iht}, \\
&\beta_h = \beta_0 + u_h, \\
&\alpha_h = \alpha_0 + v_h, \\
&e_{iht} \overset{iid}{\sim} N(0, \sigma_e^2), u_h \overset{iid}{\sim} N(0, \sigma_u^2), v_h \overset{iid}{\sim} N(0, \sigma_v^2),
\end{aligned} \tag{1}$$

where $h = 1, 2, \cdots, H, t = 1, 2, \cdots, T, i = 1, 2, \cdots, N_{ht}$ and N_{ht} is the total number of episodes in the h-th hospital and the t-th year, X_{iht} is a binary variable measuring the level of resources consumption during an episode, and $X_{iht} = 1$ if it is a hip replacement with major complexity and $X_{iht} = 0$ if with minor complexity. To make patient heterogeneity not confound the length of stay analysis, we use this as a measure of the sickness. The specification of slope β_h allows for the varying effect between LOS and severity across hospitals. Z_{iht} is a list of explanatory variables including both patient and hospital characteristics. α_h consists of two components: the common intercept α_0 applying to all episodes and the unobserved hospital effects v_h that is independently identically distributed for all h. The error terms e_{iht} embodies the remaining unexplained variations across hospitals and over time which is assumed to i.i.d. follow a normal distribution.

By introducing proper stochastic specifications, we let β_h and α_h be random variables. The specification of such random coefficients substantially reduces the number of parameters to be estimated but still allows the coefficients to differ across different hospitals. It makes the model parsimonious hence reliable estimates of β_h and α_h are easily obtained. Moreover, it is easy to draw inference about the population.

3.2 Likelihood

As the error component follows a normal distribution with its mean being zero and variance σ_e^2, the density of y_{iht} is

$$\begin{aligned} & p(y_{iht}|\beta_h,\gamma,\alpha_h,\delta_1,\delta_2,\sigma_e^2) \\ =& \left(\frac{1}{\sqrt{2\pi\sigma_e^2}}\right)\exp\left\{-\frac{1}{2\sigma_e^2}\left[y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\right]^2\right\}. \end{aligned} \tag{2}$$

Let $\boldsymbol{y}$ be a vector of all observations of y_{iht}, and its joint density is

$$\begin{aligned} & p(\boldsymbol{y}|\beta,\gamma,\alpha,\delta_1,\delta_2,\sigma_e^2) \\ =& \left(\frac{1}{\sqrt{2\pi\sigma_e^2}}\right)^N \exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\left[y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\right]^2\right\}, \end{aligned}$$

where $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_H)^\top$ and $\beta=(\beta_1,\beta_2,\cdots,\beta_H)^\top$. As α and β are assumed to be random in the panel data model given by (1), we have $\beta\sim N(\beta_0,\sigma_u^2 I_H)$ and $\alpha\sim N(\alpha_0,\sigma_v^2 I_H)$, where I_H represents the $H\times H$ identity matrix.

Therefore, the likelihood of $\boldsymbol{y}$ is

$$\begin{aligned} & p(\boldsymbol{y}|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2) = p(\boldsymbol{y}|\beta,\gamma,\alpha,\delta_1,\delta_2,\sigma_e^2)\times p(\beta|\beta_0,\sigma_u^2)\times p(\alpha|\alpha_0,\sigma_v^2) \\ \propto& (\sigma_e^2)^{-\frac{N}{2}}(\sigma_u^2)^{-\frac{H}{2}}(\sigma_v^2)^{-\frac{H}{2}}\exp\left\{-\frac{1}{2\sigma_u^2}\sum_{h=1}^{H}(\beta_h-\beta_0)^2\right\}\exp\left\{-\frac{1}{2\sigma_v^2}\sum_{h=1}^{H}(\alpha_h-\alpha_0)^2\right\} \\ & \times\exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\left[y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\right]^2\right\}. \end{aligned} \tag{3}$$

3.3 Priors

We explore Bayesian estimation of the panel data model given by (1). We choose non-informative priors for γ, δ_1 and δ_2, which are denoted as

$$p(\gamma)\propto 1, \qquad p(\delta_1)\propto 1, \qquad p(\delta_2)\propto 1.$$

The priors of β_0 and α_0 are respectively, chosen to be the densities of $N(\mu_1,\sigma_1^2)$ and $N(\mu_2,\sigma_2^2)$ with μ_1,σ_1^2,μ_2 and σ_2^2 being hyperparameters. The prior of σ_e^2 is assumed to be the inverse Gamma density denoted as $\mathrm{IG}(a_e, b_e)$ whose density is given by

$$p(\sigma_e^2) = \left(\sigma_e^2\right)^{-a_e-1} \exp\left\{-\frac{b_e}{\sigma_e^2}\right\}. \tag{4}$$

The priors of σ_u^2 and σ_v^2 are chosen to be $\mathrm{IG}(a_u, b_u)$ and $\mathrm{IG}(a_v, b_v)$, respectively. Note that a_e, b_e, a_u, b_u, a_v and b_v are hyperparameters.

The joint prior of all parameters denoted as $p(\gamma, \delta_1, \delta_2, \beta_0, \alpha_0, \sigma_e^2, \sigma_u^2, \sigma_v^2)$ is the product of the above marginal priors.

3.4 Posterior

According to Bayes theorem, the posterior of all parameters are proportional to the product of the likelihood and the joint prior:

$$\begin{aligned}
&\pi(\gamma, \delta_1, \delta_2, \beta_0, \alpha_0, \sigma_e^2, \sigma_u^2, \sigma_v^2, \beta, \alpha | y) \\
\propto& p(y|\gamma, \delta_1, \delta_2, \beta_0, \alpha_0, \sigma_e^2, \sigma_u^2, \sigma_v^2) \times p(\gamma, \delta_1, \delta_2, \beta_0, \alpha_0, \sigma_e^2, \sigma_u^2, \sigma_v^2) \\
=& \exp\left\{-\frac{1}{\sigma_e^2}\left[\frac{1}{2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\left(y_{iht} - X_{iht}\beta_h - Z_{iht}^{\top}\gamma - \alpha_h - \delta_1 t - (\delta_2 - \delta_1) t I(t \geq 7)\right)^2 + b_e\right]\right\} \\
&\exp\left\{-\frac{1}{\sigma_u^2}\left[\frac{1}{2}\sum_{h=1}^{H}(\beta_h - \beta_0)^2 + b_u\right]\right\} \exp\left\{-\frac{1}{\sigma_v^2}\left[\frac{1}{2}\sum_{h=1}^{H}(\alpha_h - \alpha_0)^2 + b_v\right]\right\} \times (\sigma_v^2)^{-\frac{H}{2} - a_v - 1} \\
&\times (\sigma_e^2)^{-\frac{N}{2} - a_e - 1} \times (\sigma_u^2)^{-\frac{H}{2} - a_u - 1} \times \exp\left\{-\frac{1}{2\sigma_1^2}(\beta_0 - \mu_1)^2\right\} \exp\left\{-\frac{1}{2\sigma_2^2}(\alpha_0 - \mu_2)^2\right\}.
\end{aligned} \tag{5}$$

Full conditionals are derived with details being given in the Appendix C, and the Gibbs sampling procedure is implemented to sample parameters from their respective conditional posteriors.

3.5 Analysis of variances

The widespread use of R^2 demonstrates a statistic that summarizes the goodness-of-fit of linear models. However, R^2 is rarely reported as a model summary statistic when mixed models are used due to theoretical problems and practical issues (see Liu et al., 2008; Orelien and Edwards, 2008; Snijders and Bosker, 1994; Xu, 2003). Recently, Nakagawa and Schielzeth (2013) extended R^2 to generalized mixed effect models by categorizing it into two types, marginal and conditional R^2_{GLMM}, that allows separation of the contributions of the fixed factors and random effects to explain variations in the response. Marginal R^2_{GLMM} gauges the variance explained by the fixed effects as a proportion of the sum of all the variances components, given by

$$R^2_{\mathrm{GLMM}(m)} = \frac{\sigma_f^2}{\sigma_f^2 + \sum_{l=1}^{u} \sigma_l^2 + \sigma_e^2}, \tag{6}$$

and conditional R^2_{GLMM} additionally includes the variance explained by the random effects in the numerator,

$$R^2_{\mathrm{GLMM}(c)} = \frac{\sigma_f^2 + \sum_{l=1}^{u} \sigma_l^2}{\sigma_f^2 + \sum_{l=1}^{u} \sigma_l^2 + \sigma_e^2}, \tag{7}$$

where σ_f^2 is the variance attributable to the fixed effects, σ_l^2 is the variance of the l–th of u random effects, and σ_e^2 is the residual variance. Since this method is only applicable to random intercept models, Johnson (2014) further extended it to encompass random slope models by calculating the mean random effect variance. Given constant mean, the variance of a mixture is simply the mean of the individual variances (Behboodian, 1970),

$$\overline{\sigma_l^2} = \frac{1}{N} \sum_j \sum_i \sigma_{lij}^2, \tag{8}$$

where N is the total number of observations. In the random coefficient model, we have

$$R^2_{\mathrm{GLMM}(m)} = \frac{\sigma_f^2}{\sigma_f^2 + \sum_{l=1}^{u} \overline{\sigma_l^2} + \sigma_e^2}, \tag{9}$$

and conditional R^2_{GLMM} additionally includes the variance explained by the random effects in the numerator,

$$R^2_{\mathrm{GLMM}(c)} = \frac{\sigma_f^2 + \sum_{l=1}^{u} \overline{\sigma_l^2}}{\sigma_f^2 + \sum_{l=1}^{u} \overline{\sigma_l^2} + \sigma_e^2}. \tag{10}$$

3.6 Posterior predictive distribution

Furthermore, we perform the counterfactual analysis by plotting the posterior predictive distributions. Suppose we have an "averaging episode" whose patient characteristics are the average of all the episodes in the dataset. At each step in the MCMC chain, we use the estimated parameters to generate a datum $\tilde{y}$ given both the averaged patient characteristics and individual hospital characteristics. Then examine the resulting distribution of $\tilde{y}$ values, which is the posterior distribution of the predicted LOS. We aim to identify any systematic differences among hospitals by doing so.

4 Results and Discussions

To assess how well the Markov chains mix, we use the simulation inefficiency factors (SIFs) of the sampler, which are calculated as the variance of the sample mean from the MCMC sampling scheme divided by the variance of the sample mean from a hypothetical sampler that draws independent random variables from the posterior (Kim et al., 1998). Considered to be a useful diagnostic measure, the calculated SIFs reported in Table D.1 are relatively low, suggesting that the proposed sampler is quite efficient in terms of producing posterior draws that are not highly correlated. Chains are wandering through the same region of the parameter space in the traceplots in Figure D.1. The autocorrelation in the corresponding correlograms drops precipitously, which means that each sample in the chain is not correlated with the previous draw and suggests no reason for concern. Lastly, the density plot allows assessment of the posterior distribution of the parameters of interest. All figures indicate that the samples are approximate independent and random draws from a parameter distribution. In the following subsections, we will focus on the effects of patient and hospital level characteristics on LOS and dig into any systematic differences between public and private hospitals by benchmarking and ranking hospitals.

4.1 Observed hospital and patient level characteristics

4.1.1 Patient demographical characteristics

The estimated coefficients of patient and hospital characteristics are given in Table 3. Age, sex and aetiology of arthritis are widely considered as the most important patient characteristics associated with the success of hip replacement. Older people experience a longer LOS due to chronic immobilization, cognitive impairment (Bo et al., 2016), increasing risk of delirium (Mahomed et al., 2005) and more medical complications postoperative. Poor health conditions and high prevalence of geriatric syndromes are extremely common among older medical inpatients.

When it comes to the causes of hip fractures, osteoporosis is one of the major risk factors. The reduced bone density will increase the risk of hip fractures, especially for the older population (Crawford and Murray, 1997). Due to a decrease in estrogen levels that occurs at female menopause and low bone density, osteoporosis is generally thought of as a "woman's disease" and leads to greater incidence of fracture than men (Cawthon, 2011). On the other hand, female patients less than 55 years old with rheumatoid arthritis are another higher risk group (Crawford and Murray, 1997). For these reasons the ALOS for females are longer than males. In general we would expect the older the patients, the more complications post surgery there are. Apart from the factors that affect LOS stated above, it is expected that the sicker patients, who have been to ICU and/or CCU, or those whose DRG codes sit in I03A (hip replacement with major complexity) will have a longer LOS. Moreover, the fact that the unmarried patients have longer LOS may reflect a lack of social support among them.

Studies of hospital LOS are incomplete unless they also include post-hospital care (Picone et al., 2003). Knowing the decision of destination to discharge a patient is a factor that significantly affect the LOS, we find LOS shorter for patients who were relocated to the post-hospital care alternatives such as nursing homes and rehabilitation centers than those who were discharged home. This could explain the a reduced LOS in the acute care hospital followed by the immediate admission to the on-site inpatient rehabilitation facility. Some patients considered to be good candidates of rehabilitation may be sent early to rehabilitation centers even though they can not possibly walk without proper assistance (Pablo et al., 2004). This phenomenon has been widely discussed and is believed to impose a heavy burden to the post-hospital care facilities (Balen et al., 2002; Weingarten et al., 1998).

Table 3: Effects of observed hospital and patient level characteristics on LOS

	Posterior Mean	Marginal Effect	SD	Credible Interval 2.50%	97.50%
constant	1.739	14.160	0.034	1.670	1.807
age	0.005	0.045	0.000	0.005	0.006
male	-0.020	-0.164	0.009	-0.037	-0.003
married	-0.015	-0.123	0.009	-0.032	0.002
int_care	0.243	2.240	0.016	0.212	0.274
emergency	0.298	2.830	0.012	0.275	0.323
pub_pub	-0.189	-1.399	0.013	-0.214	-0.163
pri_pub	-0.215	-1.574	0.020	-0.255	-0.175
sep_N	-0.098	-0.757	0.029	-0.155	-0.039
sep_S	-0.079	-0.616	0.014	-0.106	-0.050
sep_T	0.004	0.031	0.010	-0.017	0.024
metro	0.206	1.860	0.016	0.174	0.238
major	0.075	0.633	0.017	0.042	0.107
δ_1	-0.021	-0.168	0.004	-0.029	-0.013
δ_2	-0.056	-0.459	0.002	-0.060	-0.053

Kelly et al. (2012) found emergency patients waited longer for surgery than elective patients and stayed longer post surgery. We confirm this by showing that patients admitted from the emergency departments stay 2.8 days longer than than those from the elective surgery waiting list. When patients spend a long time in emergency department, their treatment won't typically get started until they actually get to a hospital floor. Therefore the delays in treatment might lead to longer hospitalization overall. Another reason is that the long queue waiting in the emergency department proposes a heavy burden on the staff. Thus the management process in emergency department became more complicated due to rapid development. Patients stay in emergency departments longer, which in return make the emergency departments more crowded. The adverse impact of emergency department crowding includes prolonged waiting times, increased complications and increased mortality (Chaou et al., 2016). Moreover, elder people are more likely to be admitted though the emergency department, compared with the young ones. The older and sicker emergency patients may have had symptoms but delayed seeing a doctor due to any possible reasons, such as their own denial and fear or the difficulties to access the hospital resources for proper treatment. The increasing proportion of emergency admissions, coupled with the evidence of increased LOS has become a recognized issue in public health system, not only for Australia (Kelly et al., 2012). These cases are generally more complicated and harder to deal with than elective patients.

We also notice a structural break in year 2009 in the data. The trend can be influenced by many factors such as the pressure to reduce the waiting time, different ways to reimburse the hospitals or changes in technology. As is mentioned previously, the decrease in LOS is more significant post 2009. The hospitals have been trying to improve their efficiency during that period and start to use minimal invasive (MI) surgery approaches in total hip replacement (Röttger et al., 2012), that were developed on the basis of conventional approaches and are supposed to reduce the LOS (Davies and Welch, 2010). Meanwhile, the total number of hip replacement has increased drastically from 2009 onwards. This brings our attention to the defective hip replacement scandal in 2009 that affected thousands of patients wellbeing. DePuy Orthopaedics, the maker of articular surface replacement (ASR) hip, issued a worldwide recall of the hip in 2010, after they discontinued the supply of the implants in Australia in 2009. Unfortunately 5500 Australians have been implanted with the faulty hip and a great number of them have suffered terribly from the side effects of the hip failing, which left a permanent damage from the heavy metal that entered their bloodstream. Patients had to revisit the hospitals to replace the faulty implants, which leads to the increase in the number of hip replacements from 2009.

4.1.2 Systematic difference in private and public hospitals

Hospital type is indicative of different administrative arrangement and budget constraints therefore more attention should be paid to coefficients of pub_pub and pri_pub in Table 3. Systematic differences in the LOS among patients in different hospitals are observed. Public patients or patients who elect private status in public hospitals stay around one and a half days shorter than privately insured patients in private hospitals. It is clear that public hospitals rush patients out of the door too quickly especially when a high proportion of patients are from disadvantaged socioeconomic areas or those who acquired relatively more complexed conditions. We know that good hospital care keeps patients illnesses under control. When conditions are not well managed, the patient's health deteriorates and they're more likely to require more costly hospital care in the future.

This phenomenon is commonly seen under ABF. Payment of this funding scheme is set by the government, based on the average cost of similar patients nationally, irrespective of where the hospital is and whether the hospital delivers care in an efficient manner. Under this fee for service payment model, the hospitals will be paid a fixed price for each episode of care. As a matter of course, it introduces the incentives for public hospitals to discharge the patients as early as they can to free up more beds, therefore maximize the reimbursement. When the cost of the episode is below the fixed price paid, hospitals make a profit. However, the most profitable service does not necessarily mean the greatest health gains for the population. This has been confirmed by the increasing readmission rate since the introduction of ABF. The policy impasse has been confirmed by many other researches showing that some hospitals have been mismanaged on an epic scale. Even though the government wants the hospitals to be more efficient under this scheme, the flaws in the scheme could hamper the potential cost savings and lead to an ineffective funding scheme. Under such circumstances, taxpayers should be worried whether the dollars have been channeled into essential hospital care or diverted on useless bureaucracy at the expense of patient wellbeing.

The systematic difference also indicate the potential moral hazard effect amongst patient in private hospitals. Private hospital insurance can not only cause an increase in the likelihood of hospital admission (see Eldridge et al., 2017) but also increase the patient LOS. Cameron et al. (1988) observed a longer LOS for privately insured patients relative to non-insured individuals. Similar conclusion was given in Savage and Wright (2003) where insurance increased the expected LOS by a factor of at least 2. Insurance has a large effect on the LOS and we will provide a detailed comparison by ranking the hospitals in later sections.

4.2 Bench marking and ranking hospitals

4.2.1 Variance partition coefficients

Equally important are to what extent the variations in LOS can be explained by unobserved hospital effects. Table 4 represents the estimated variance of the hospital level and random error components, percentage contribution of observable patient and hospital characteristics to total variations in the LOS and the proportion of the unexplained variation attributable to the random hospital effect. The result for marginal and conditional R^2 shows that 31.6 percent of the variation in LOS are explained by the observable characteristics and 5.6 percent can be explained be the random components. The remaining 60.9 percent unexplained variance can be decomposed into the within and between hospital variances. After controlling for hospital and patient characteristics, a large proportion of the variation comes from the patients themselves and 61.4 percent of the unexplained variation is due to the unknown features of the hospitals where they were treated. This gives an indication of the size of the differences in LOS attributable to hospital level variation, which can be compared with the differences attributed to patient level covariates. Possible explanations are given in the following subsections.

Table 4: Estimated variances and variance partitions

σ_e^2	0.197
σ_v^2	0.016
σ_u^2	0.010
Variation partition (%): observable factors and unobservable factors	
$R^2_{GLMM}(m)$: % explained by fixed effects	31.6
$R^2_{GLMM}(c)$: % explained by both fixed and random effects	37.2

4.2.2 Individual hospital effects

The common effect of complexity β_0 is estimated to be 0.33, which means hip replacement patients with major complexity stay 3 days longer than those with minor complexity. However we believe that the effect of patient complexity levels on LOS is not constant across hospitals. For example, specialty hospitals may be more efficient than full-service hospitals as they may adopt different technologies, doctors and nurses may have different skills and hospital size may affect the hospital output when delivering services to patients with varied complexity. It is therefore essential to differentiate variations in LOS caused by major or minor complexity from individual hospitals. Results in Table 5 confirmed our hypothesis with the variations presented. As the public hospitals treat a disproportionately large share of people of low socioeconomic status, patients are likely to be characterised by a relatively poor pre-existing health status and relatively more health-related risk factors. Taking a closer look at the marginal effect, patients with major complextity (DRG code I03A) in hospital 15 stay 5 days longer than those with minor complexity (DRG code I03B), while the gap drops to 1 day in hospital 11. It may be due to the fact that hospital 15 is a major hospital and the patients there are sicker than hospital 11 even though they are categorized by the same DRG code. The difference could also come from the location of the hospitals. Hospitals located in remote areas are likely to incur a higher cost of transporting hospital supplies as well as greater difficulty attracting staff, e.g., hospital 15. In general, many regional hospitals are very small in capacity and therefore unable to benefit from economies of scale.

Even among similar hospitals, the differences in LOS are not always captured by the performance data. We tackle this problem by introducing the unobserved hospital effects v_h in the model. Australian government has been under increasing pressure to improve the healthcare quality while seeking to effectively employ their scarce resources. Huge disparities still exist in the cost of delivering the same service across different hospitals. A common hip replacement can cost \$16,000 more in the most expensive hospital compared with the cheapest. When the finances are tight, the large differences cry out for attention. Table 6 presents the posterior mean, standard deviation and 95% credible interval of each unobserved hospital effect. h_1 to h_{20} are 20 public hospitals and p_1 to p_{20} represent 20 private hospitals. Total rank is the ranking of all the hospitals whereas group rank is the ranking of hospitals within public or private groups separately. Magnitudes and 95% credible intervals of v_h indicate very different distributions of the unobserved effects. As we can see, v_h in major providers and teaching hospitals are higher than others, which can be explained partially by the higher cost of labor, more complex case mix (Freitas et al., 2012) and greater obligation of medical graduate education (Omachonu et al., 2004).

To better interpret the posterior density distributions, we plot the distribution of all posteriors in Figure 2. Some public hospitals are significantly outperformed by others suggesting room to reduce the avoidable cost and to improve the performance. Among all public hospitals, the highest v_h is centering around 0.17, with the smallest being -0.22. The corresponding results are 0.36 and -0.27 for private hospitals. Overall, the distribution of public hospitals are tightened up due to a more rapid diffusion of knowledge on medical procedures and medical treatment across providers and pressure from the government. A larger variation is clearly spotted among private ones and management level should be given information about their relative performance so they can identify potential saving opportunities. Hospitals should also investigate whether the variations come from the department or is it due to the supplies or staff as rigidities in the roles that doctors and nurses perform also add to costs. Different hospitals are facing different problems that require different solutions. The Department of Health and Human Services (DHHS) did not provide sufficient overarching framework or single repository of information to benchmark hospital performance. Hence rooting out variations in hospitals is a good way to start making healthcare affordable and keeping health budget in control.

Table 5: Estimated varying beta

	Posterior Mean	Marginal Effect	SD	Credible Interval 2.50%	Credible Interval 97.50%	Total Rank	Metro	Major
β_1	0.35	3.38	0.02	0.30	0.40	23	1	1
β_2	0.50	5.22	0.02	0.45	0.54	37	1	1
β_3	0.18	1.61	0.03	0.13	0.23	4	1	1
β_4	0.29	2.75	0.03	0.24	0.34	17	1	1
β_5	0.29	2.74	0.02	0.25	0.33	16	1	1
β_6	0.32	3.09	0.03	0.26	0.38	19	1	0
β_7	0.32	3.10	0.03	0.27	0.38	20	1	1
β_8	0.19	1.73	0.04	0.12	0.27	6	1	0
β_9	0.20	1.85	0.04	0.13	0.28	7	1	0
β_{10}	0.43	4.39	0.03	0.38	0.48	32	1	1
β_{11}	0.14	1.26	0.06	0.02	0.27	2	1	0
β_{12}	0.44	4.59	0.03	0.39	0.50	34	1	1
β_{13}	0.22	1.99	0.03	0.17	0.27	8	0	1
β_{14}	0.19	1.70	0.03	0.12	0.26	5	0	1
β_{15}	0.50	5.27	0.02	0.46	0.54	38	0	1
β_{16}	0.35	3.39	0.04	0.27	0.43	24	0	0
β_{17}	0.36	3.55	0.03	0.30	0.43	27	0	0
β_{18}	0.28	2.60	0.03	0.22	0.34	14	0	1
β_{19}	0.33	2.21	0.02	0.29	0.38	21	1	1
β_{20}	0.48	5.00	0.03	0.41	0.55	36	0	1
β_{21}	0.39	3.84	0.05	0.29	0.49	28	1	0
β_{22}	0.34	3.32	0.02	0.30	0.38	22	1	1
β_{23}	0.43	4.42	0.03	0.38	0.49	33	1	1
β_{24}	0.43	4.38	0.05	0.34	0.52	31	1	0
β_{25}	0.30	2.89	0.07	0.16	0.45	18	0	0
β_{26}	0.24	2.26	0.05	0.15	0.34	11	0	0
β_{27}	0.56	6.17	0.04	0.48	0.65	40	0	1
β_{28}	0.29	2.72	0.05	0.19	0.38	15	0	1
β_{29}	0.26	2.38	0.05	0.15	0.36	12	1	0
β_{30}	0.24	2.18	0.05	0.13	0.34	9	1	0
β_{31}	0.42	4.26	0.06	0.31	0.54	29	1	0
β_{32}	0.52	5.50	0.04	0.43	0.60	39	1	1
β_{33}	0.24	2.21	0.04	0.17	0.31	10	1	1
β_{34}	0.47	4.93	0.03	0.41	0.54	35	1	1
β_{35}	0.43	4.33	0.04	0.35	0.50	30	1	1
β_{36}	0.36	3.48	0.06	0.23	0.48	26	1	1
β_{37}	0.17	1.55	0.07	0.03	0.31	3	0	0
β_{38}	0.35	3.40	0.07	0.21	0.49	25	1	0
β_{39}	0.27	2.53	0.07	0.14	0.40	13	0	0
β_{40}	0.10	0.86	0.03	0.04	0.16	1	1	1

Table 6: Hospital performance: posterior summary statistics for v_h

	Posterior Mean	SD	Credible Interval 2.50%	97.50%	Total Rank	Group Rank	Metro	Major
h_1	-0.09	0.04	-0.16	-0.02	12	8	1	1
h_2	-0.16	0.04	-0.23	-0.09	4	2	1	1
h_3	-0.11	0.04	-0.18	-0.04	8	5	1	1
h_4	0.08	0.04	0.01	0.15	28	17	1	1
h_5	-0.10	0.04	-0.17	-0.03	11	7	1	1
h_6	-0.09	0.04	-0.17	-0.01	13	9	1	0
h_7	-0.03	0.04	-0.10	0.04	22	14	1	1
h_8	-0.04	0.04	-0.12	0.03	20	13	1	0
h_9	0.04	0.04	-0.04	0.11	26	16	1	0
h_{10}	-0.07	0.04	-0.14	0.00	16	12	1	1
h_{11}	-0.14	0.04	-0.22	-0.06	6	3	1	0
h_{12}	0.11	0.03	0.04	0.18	30	18	1	1
h_{13}	-0.08	0.03	-0.14	-0.01	15	11	0	1
h_{14}	-0.11	0.03	-0.17	-0.04	9	6	0	1
h_{15}	0.17	0.03	0.11	0.24	36	20	0	1
h_{16}	-0.22	0.03	-0.29	-0.16	2	1	0	0
h_{17}	0.01	0.03	-0.05	0.07	24	15	0	0
h_{18}	-0.08	0.03	-0.15	-0.02	14	10	0	1
h_{19}	-0.14	0.04	-0.21	-0.06	7	4	1	1
h_{20}	0.14	0.03	0.08	0.20	33	19	0	1
p_1	0.14	0.03	0.07	0.20	32	14	1	0
p_2	-0.03	0.03	-0.09	0.03	21	8	1	1
p_3	-0.06	0.03	-0.12	0.01	17	5	1	1
p_4	0.04	0.03	-0.03	0.10	27	11	1	0
p_5	0.22	0.03	0.16	0.29	37	17	0	0
p_6	0.36	0.03	0.30	0.42	40	20	0	0
p_7	-0.27	0.03	-0.33	-0.21	1	1	0	1
p_8	0.14	0.03	0.08	0.20	34	15	0	1
p_9	0.11	0.03	0.05	0.18	31	13	1	0
p_{10}	0.15	0.03	0.09	0.21	35	16	1	0
p_{11}	0.09	0.03	0.02	0.15	29	12	1	0
p_{12}	-0.22	0.03	-0.28	-0.15	3	2	1	1
p_{13}	-0.05	0.03	-0.11	0.01	18	6	1	1
p_{14}	-0.04	0.03	-0.10	0.02	19	7	1	1
p_{15}	-0.15	0.03	-0.21	-0.09	5	3	1	1
p_{16}	-0.02	0.03	-0.09	0.04	23	9	1	1
p_{17}	0.29	0.03	0.23	0.36	38	18	0	0
p_{18}	0.01	0.03	-0.05	0.08	25	10	1	0
p_{19}	0.34	0.03	0.28	0.41	39	19	0	0
p_{20}	-0.10	0.03	-0.16	-0.04	10	4	1	1

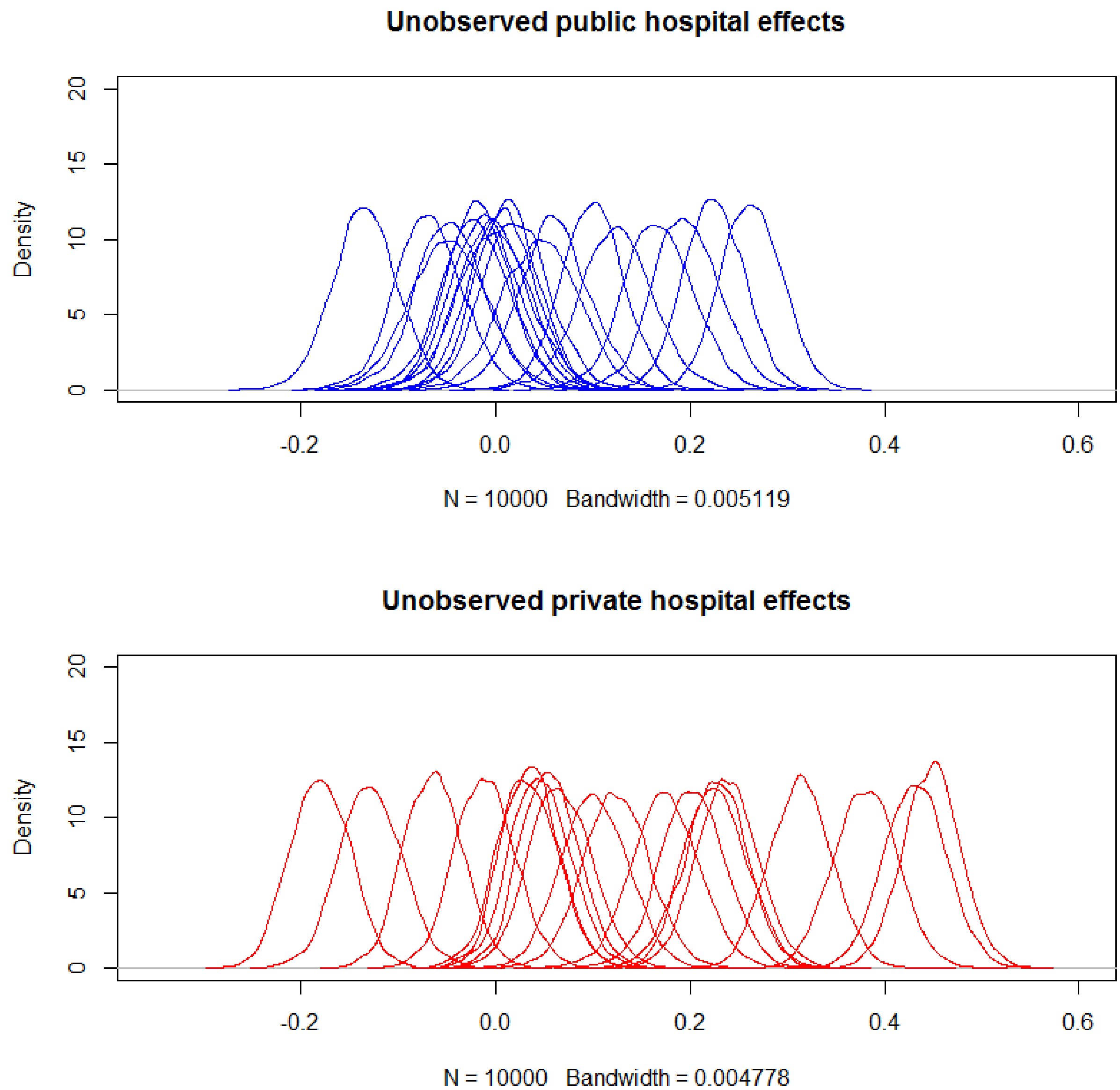


Figure 2: Unobserved hospital effects for all hospitals

4.2.3 Counterfactual analysis

We further plot the posterior predictive distributions for LOS in Figure 3. Hypothetically giving all hospitals the same patients (an average patient) in year 2014, the differences among hospitals mainly come from three sources - hospital treatment of patients with different levels of complexity, unobserved hospital effects and size and location of the hospitals. It is clear in Table 7 that the predictive posterior mean for most public hospitals range from 3 to 6 days but in private hospitals the interval extends to 5 to 7 days. We can clearly see some patterns from the behavior of individual hospitals. For example, hospitals with the longest predicted LOS are all private such as p_1, p_6, p_{10} and p_{19}. The number of hip replacement performed is low compared with other hospitals. These hospitals do not necessarily have the bed crisis as the large general hospitals and specialty hospitals do[4]. On top of this, patient privately insured may be allowed to stay relatively long as the cost of stay is covered by the insurance companies. On the contrary, almost all the public hospitals have a shorter predicted LOS, regardless of whether they are located in rural or metropolitan areas, e.g. h_{16} and h_{14}.

ALOS are 4.6 and 6 days for public and private hospitals respectively[5]. The 1.4 days difference in ALOS is significant as the ALOS is already extremely short for both types of hospitals. Patients in private hospitals stay 30 percent longer than those in the public hospitals. This is a striking finding that leads up to the discussion of 'quicker but sicker' phenomenon among public hospitals in recent years. Although shortening hospital stays is a worthwhile policy objective, patients might require readmission to hospitals too soon if discharged prematurely. Almost certainly, many will need some extent of post-acute care. It is a moral imperative not to shift cost to the patients and it is certainly unethical to put them into a situation where their safety and quality of care are jeopardized, particularly when hospitals can use the resources more appropriately and reduce unnecessary cost with better management. Resolving these contentious issues in our health system will take time therefore the hospitals should be aiming to improve the performance by increasing the success rate of medical treatment and surgical procedures, instead of playing games on the funding scheme. It is also potentially harmful to the whole healthcare system as the revisit of patients will result in longer waiting time and heavier burden to the rehabilitation centers, nursing homes and aged care.

In the long run, a specialty that keeps functioning at a loss can be completely shut down by the hospital so that they could use the resources in other specialties to make a profit. It is also likely that the hospital will 'upcode' the DRG hoping to get more funding for the same patient. There is also possibility that they break one episode into several to

[4]We suspect that the reason why p_7 stands out with a significantly short predicted LOS is that it attracts many patients from surrounding rural areas and exhibit a high volume of patients, hence lots of pressure to free up beds as soon as they can.

[5]We also performed the counterfactual analysis using year 2003, the ALOS is 8.8 days for public hospitals and 11.6 days for private hospitals.

Table 7: Posterior predictive distribution for LOS in 2014

	Posterior Mean	SD	Credible Interval 2.50%	Credible Interval 97.50%	Total Rank	Group Rank	Metro	Major
h_1	4.77	0.07	4.63	4.91	14	13	1	1
h_2	4.57	0.06	4.45	4.69	12	11	1	1
h_3	4.56	0.07	4.42	4.70	11	10	1	1
h_4	5.60	0.09	5.42	5.78	23	19	1	1
h_5	4.69	0.07	4.55	4.82	13	12	1	1
h_6	4.41	0.09	4.25	4.58	8	7	1	0
h_7	5.06	0.08	4.90	5.22	18	17	1	1
h_8	4.53	0.11	4.32	4.75	9	8	1	0
h_9	4.91	0.10	4.71	5.11	15	14	1	0
h_{10}	4.95	0.07	4.80	5.09	16	15	1	1
h_{11}	4.09	0.10	3.89	4.30	7	6	1	0
h_{12}	5.91	0.09	5.73	6.09	26	20	1	1
h_{13}	3.86	0.09	3.69	4.03	3	3	0	1
h_{14}	3.73	0.08	3.58	3.88	2	2	0	1
h_{15}	5.19	0.11	4.97	5.41	19	18	0	1
h_{16}	3.16	0.08	3.00	3.32	1	1	0	0
h_{17}	4.00	0.10	3.80	4.20	5	5	0	0
h_{18}	3.87	0.09	3.70	4.05	4	4	0	1
h_{19}	4.55	0.08	4.40	4.71	10	9	1	1
h_{20}	4.97	0.09	4.80	5.16	17	16	0	1
p_1	6.77	0.18	6.43	7.12	40	20	1	0
p_2	6.12	0.08	5.96	6.29	30	10	1	1
p_3	6.07	0.09	5.90	6.24	29	9	1	1
p_4	6.17	0.15	5.88	6.48	31	11	1	0
p_5	5.95	0.15	5.65	6.26	27	7	0	0
p_6	6.74	0.14	6.48	7.02	39	19	0	0
p_7	4.08	0.07	3.95	4.21	6	1	0	1
p_8	5.87	0.10	5.67	6.07	24	5	0	1
p_9	6.49	0.14	6.22	6.77	36	16	1	0
p_{10}	6.70	0.13	6.44	6.97	38	18	1	0
p_{11}	6.48	0.15	6.19	6.78	35	15	1	0
p_{12}	5.24	0.08	5.07	5.40	20	2	1	1
p_{13}	5.90	0.07	5.76	6.05	25	6	1	1
p_{14}	6.19	0.08	6.03	6.35	32	12	1	1
p_{15}	5.51	0.07	5.36	5.65	21	3	1	1
p_{16}	6.20	0.11	5.99	6.42	33	13	1	1
p_{17}	6.23	0.16	5.92	6.55	34	14	0	0
p_{18}	5.96	0.14	5.69	6.24	28	8	1	0
p_{19}	6.66	0.17	6.34	6.99	37	17	0	0
p_{20}	5.51	0.08	5.36	5.66	22	4	1	1

maximize the amount of money they can possibly get. This vicious cycle leads to patient dumping, cream skimming and longer waiting times for hip replacement and reduced efficiency in the whole system and the consequence is shared by both the government and patients.

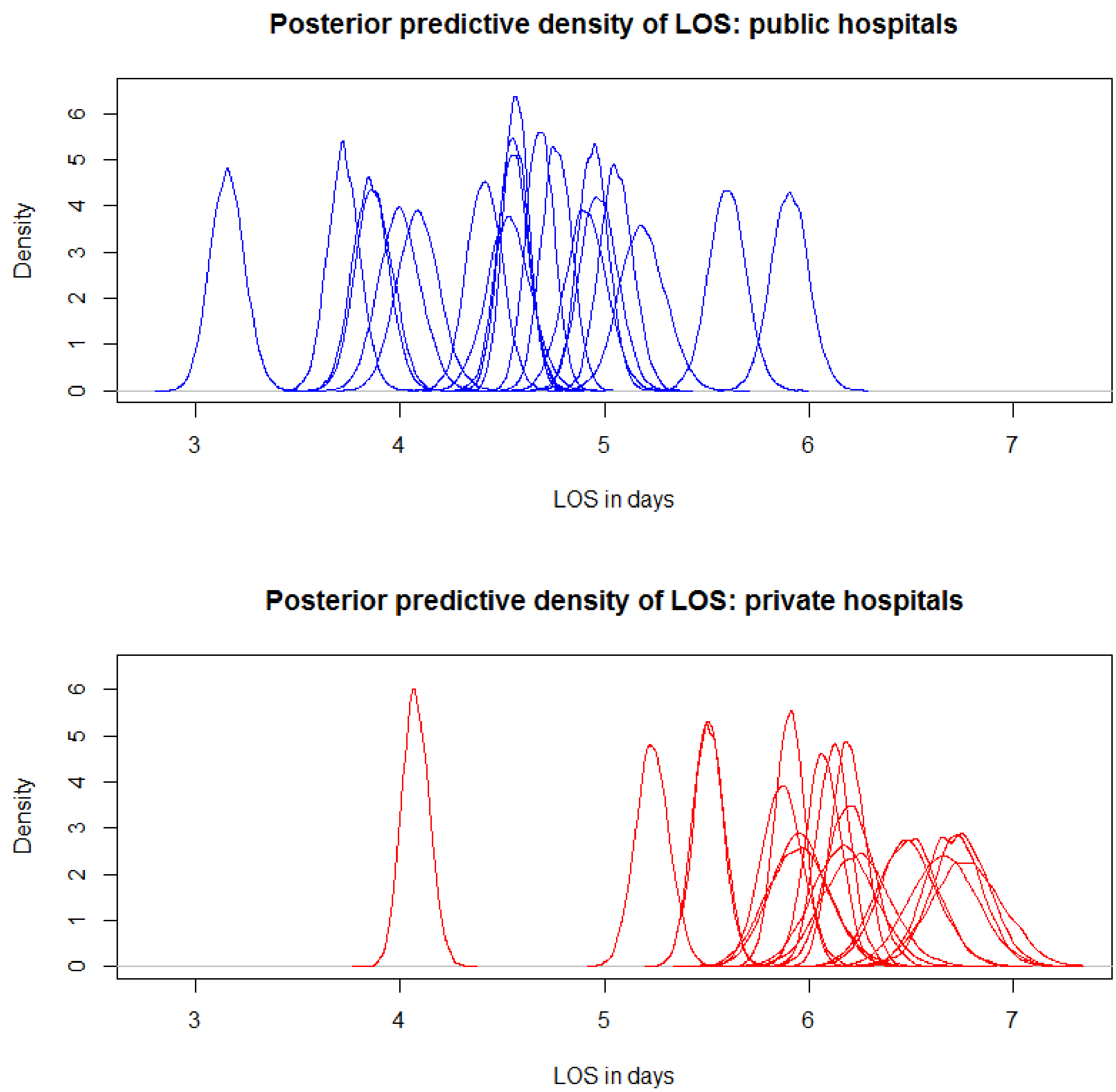


Figure 3: Posterior predictive distribution for all hospitals

5 Conclusion

The paper investigates the variations in hospital LOS of hip replacement patients using a large scale admitted hospital dataset of Victoria. We estimate a Bayesian hierarchical random coefficient model to study the observed patient and hospital characteristics and the unobserved hospital effect on LOS. The work demonstrates the advantages of Bayesian approach when dealing with large multilevel data, adds to current knowledge of factors that affect LOS and brings the attention of the public to build a healthier healthcare system.

The introduction of ABF forces hospitals to find the most cost-effective means of delivering care and be more responsive to patient demand but it is certainly not far-reaching enough because it does not address the dysfunctional mismanagement of the hospitals. We find that widespread variation in LOS exist in Victorian hospitals, which means some failed to make full use of the beds available and therefore increased the unnecessary cost. Variation in private hospitals are significantly higher than public hospitals, requiring management levels in these hospitals to investigate the causes and to reduce the cost in order to catch up with the peers. Policy makers should support such moves by making the data available to hospitals for comparison, release measurement reports in shorter cycles and investigate the worst performing hospitals carefully.

Apart from the unsatisfactory performance of some hospitals, we also confirm the 'quicker but sicker' phenomenon. In order to increase the marginal revenue they receive from the government, public hospitals rush patients out of the door for 1.4 days earlier than the private ones given a patient with same level of complexity, in spite of shifting tremendous risks to the patients. Reducing the LOS is a worthwhile policy objective only on the condition that the quality of care is maintained. The challenge hospitals and states are facing is great. Meeting them will require sophisticated strategies and innovation and there is a long journey ahead for the funding policy to achieve successes. But early efforts deserve more than knee-jerk criticism based on perceptions formed decades ago about what hospitals can do. Interventions to fix the problems will have to be multidimensional and involve the efforts of the government, hospitals and patients.

At this point, some caveats should be reinforced. We do not have full control of hospital characteristics, such as the teaching status of the private hospitals and locations of the hospitals. In terms of the unobserved characteristics, we have no information on the department specific factors such as the safety procedures and protocols that affect the patient complexity levels. Despite these limitations, the result of our research will contribute evaluating not only the efficiency but also the the effectiveness of ABF policy, motivating hospitals to better allocate their resources and most importantly, improve the patients' experience of care.

References

Adams, G., Gulliford, M., Ukoumunne, O., Chinn, S. and Campbell, M. (2003), 'Geographical and organisational variation in the structure of primary care services: implications for study design', *Journal of Health Services Research & Policy* **8**(2), 87–93.

AIHW (2012), *Surgery in Australian hospitals*, Australian Institute of Health and Welfare.

Austin, P. C., Rothwell, D. M. and Tu, J. V. (2002), 'A comparison of statistical modeling strategies for analyzing length of stay after cabg surgery', *Health Services and Outcomes Research Methodology* **3**(2), 107–133.

Balen, R. V., Steyerberg, E. W., Cools, H. J., Polder, J. J. and Habbema, J. D. F. (2002), 'Early discharge of hip fracture patients from hospital: transfer of costs from hospital to nursing home', *Acta Orthopaedica Scandinavica* **73**(5), 491–501.

Behboodian, J. (1970), 'On the modes of a mixture of two normal distributions', *Technometrics* **12**(1), 131–139.

Bo, M., Fonte, G., Pivaro, F., Bonetto, M., Comi, C., Giorgis, V., Marchese, L., Isaia, G., Maggiani, G., Furno, E. et al. (2016), 'Prevalence of and factors associated with prolonged length of stay in older hospitalized medical patients.', *Geriatrics & gerontology international* **16**(3), 314–321.

Bueno, H., Ross, J. S., Wang, Y., Chen, J., Vidán, M. T., Normand, S.-L. T., Curtis, J. P., Drye, E. E., Lichtman, J. H., Keenan, P. S. et al. (2010), 'Trends in length of stay and short-term outcomes among medicare patients hospitalized for heart failure, 1993-2006', *JAMA* **303**(21), 2141–2147.

Cameron, A. C., Trivedi, P. K., Milne, F. and Piggott, J. (1988), 'A microeconometric model of the demand for health care and health insurance in Australia', *The Review of economic studies* **55**(1), 85–106.

Carey, K. (2015), 'Measuring the hospital length of stay/readmission cost trade-off under a bundled payment mechanism', *Health Economics* **24**(7), 790–802.

Castelli, A., Daidone, S., Jacobs, R., Kasteridis, P. and Street, A. D. (2015), 'The determinants of costs and length of stay for hip fracture patients', *PloS One* **10**(7), e0133545.

Castelli, A., Jacobs, R., Goddard, M. and Smith, P. C. (2013), 'Health, policy and geography: insights from a multi-level modelling approach', *Social Science & Medicine* **92**, 61–73.

Cawthon, P. M. (2011), 'Gender differences in osteoporosis and fractures', *Clinical Orthopaedics and Related Research* **469**(7), 1900–1905.

Chaou, C.-H., Chiu, T.-F., Yen, A. M.-F., Ng, C.-J. and Chen, H.-H. (2016), 'Analyzing factors affecting emergency department length of stay–using a competing risk-accelerated failure time model', *Medicine* **95**(14), e3263.

Charlson, M., Szatrowski, T. P., Peterson, J. and Gold, J. (1994), 'Validation of a combined comorbidity index', *Journal of Clinical Epidemiology* **47**(11), 1245–1251.

Cheng, T. C., Palangkaraya, A. and Yong, J. (2014), 'Hospital utilization in mixed public–private system: evidence from Australian hospital data', *Applied Economics* **46**(8), 859–870.

Correia, M. I. T. and Waitzberg, D. L. (2003), 'The impact of malnutrition on morbidity, mortality, length of hospital stay and costs evaluated through a multivariate model analysis', *Clinical Nutrition* **22**(3), 235–239.

Cram, P., Lu, X., Kaboli, P. J., Vaughan-Sarrazin, M. S., Cai, X., Wolf, B. R. and Li, Y. (2011), 'Clinical characteristics and outcomes of medicare patients undergoing total hip arthroplasty, 1991-2008', *JAMA* **305**(15), 1560–1567.

Crawford, R. and Murray, D. (1997), 'Total hip replacement: indications for surgery and risk factors for failure', *Annals of the Rheumatic Diseases* **56**(8), 455–457.

Davies, L. and Welch, H. G. (2010), 'Thyroid cancer survival in the united states: observational data from 1973 to 2005', *Archives of Otolaryngology–Head & Neck Surgery* **136**(5), 440–444.

Draper, B. and Luscombe, G. (1998), 'Quantification of factors contributing to length of stay in an acute psychogeriatrics ward', *International Journal of Geriatric Psychiatry* **13**(1), 1–7.

Eastwood, E. A., Hagglund, K. J., Ragnarsson, K. T., Gordon, W. A. and Marino, R. J. (1999), 'Medical rehabilitation length of stay and outcomes for persons with traumatic spinal cord injury–1990-1997', *Archives of Physical Medicine and Rehabilitation* **80**(11), 1457–1463.

Eldridge, D. S., Onur, I. and Velamuri, M. (2017), 'The impact of private hospital insurance on the utilization of hospital care in Australia', *Applied Economics* **49**(1), 78–95.

Freitas, A., Silva-Costa, T., Lopes, F., Garcia-Lema, I., Teixeira-Pinto, A., Brazdil, P. and Costa-Pereira, A. (2012), 'Factors influencing hospital high length of stay outliers', *BMC Health Services Research* **12**(1), 265.

Hall, J. (2010), 'Health-care reform in Australia: advancing or side-stepping?', *Health Economics* **19**(11), 1259–1263.

Hauck, K. and Hollingsworth, B. (2008), Do obese patients stay longer in hospital?: Estimating the health care costs of obesity. Monash University.

Husted, H., Hansen, H. C., Holm, G., Bach-Dal, C., Rud, K., Andersen, K. L. and Kehlet, H. (2010), 'What determines length of stay after total hip and knee arthroplasty? a nationwide study in denmark', *Archives of Orthopaedic and Trauma Surgery* **130**(2), 263.

Husted, H., Holm, G. and Jacobsen, S. (2008), 'Predictors of length of stay and patient satisfaction after hip and knee replacement surgery: fast-track experience in 712 patients', *Acta Orthopaedica* **79**(2), 168–173.

Jensen, P. H., Webster, E. and Witt, J. (2009), 'Hospital type and patient outcomes: an empirical examination using ami readmission and mortality records', *Health Economics* **18**(12), 1440–1460.

Johnson, P. C. (2014), 'Extension of nakagawa & schielzeth's r2glmm to random slopes models', *Methods in Ecology and Evolution* **5**(9), 944–946.

Jones, A. M. (2000), 'Health econometrics', *Handbook of Health Economics* **1**, 265–344.

Kelly, M., Sharp, L., Dwane, F., Kelleher, T. and Comber, H. (2012), 'Factors predicting hospital length-of-stay and readmission after colorectal resection: a population-based study of elective and emergency admissions', *BMC Health Services Research* **12**(1), 77.

Kim, S., Shephard, N. and Chib, S. (1998), 'Stochastic volatility: likelihood inference and comparison with arch models', *The Review of Economic Studies* **65**(3), 361–393.

Koop, G., Osiewalski, J. and Steel, M. F. (1997), 'Bayesian efficiency analysis through individual effects: Hospital cost frontiers', *Journal of Econometrics* **76**(1), 77–105.

Learmonth, I. D., Young, C. and Rorabeck, C. (2007), 'The operation of the century: total hip replacement', *The Lancet* **370**(9597), 1508–1519.

Librero, J., Peiró, S. and Ordiñana, R. (1999), 'Chronic comorbidity and outcomes of hospital care: length of stay, mortality, and readmission at 30 and 365 days', *Journal of Clinical Epidemiology* **52**(3), 171–179.

Liu, H., Zheng, Y. and Shen, J. (2008), 'Goodness-of-fit measures of R^2 for repeated measures mixed effect models', *Journal of Applied Statistics* **35**(10), 1081–1092.

Liu, Y., Phillips, M. and Codde, J. (2001), 'Factors influencing patients' length of stay', *Australian Health Review* **24**(2), 63–70.

Mahomed, N. N., Barrett, J., Katz, J. N., Baron, J. A., Wright, J. and Losina, E. (2005), 'Epidemiology of total knee replacement in the united states medicare population', *Journal of Bone & Joint Surgery* **87**(6), 1222–1228.

Martin, S., Street, A., Han, L. and Hutton, J. (2016), 'Have hospital readmissions increased in the face of reductions in length of stay? evidence from england', *Health Policy* **120**(1), 89–99.

Morales, I. J., Peters, S. G. and Afessa, B. (2003), 'Hospital mortality rate and length of stay in patients admitted at night to the intensive care unit', *Critical Care Medicine* **31**(3), 858–863.

Moran, J. L. and Solomon, P. J. (2012), 'A review of statistical estimators for risk-adjusted length of stay: analysis of the Australian and New Zealand intensive care adult patient data-base, 2008-2009', *BMC Medical Research Methodology* **12**(1), 68.

Nakagawa, S. and Schielzeth, H. (2013), 'A general and simple method for obtaining R^2 from generalized linear mixed-effects models', *Methods in Ecology and Evolution* **4**(2), 133–142.

Olsen, K. R. and Street, A. (2008), 'The analysis of efficiency among a small number of organisations: How inferences can be improved by exploiting patient-level data', *Health Economics* **17**(6S), 671–681.

Omachonu, V. K., Suthummanon, S., Akcin, M. and Asfour, S. (2004), 'Predicting length of stay for medicare patients at a teaching hospital', *Health Services Management Research* **17**(1), 1–12.

Orelien, J. G. and Edwards, L. J. (2008), 'Fixed-effect variable selection in linear mixed models using r2 statistics', *Computational Statistics & Data Analysis* **52**(4), 1896–1907.

Pablo, P. d., Losina, E., Phillips, C. B., Fossel, A. H., Mahomed, N., Lingard, E. A. and N Katz, J. (2004), 'Determinants of discharge destination following elective total hip replacement', *Arthritis Care & Research* **51**(6), 1009–1017.

Palangkaraya, A. and Yong, J. (2013), 'Effects of competition on hospital quality: an examination using hospital administrative data', *The European Journal of Health Economics* **14**(3), 415–429.

Picone, G., Mark Wilson, R. and Chou, S.-Y. (2003), 'Analysis of hospital length of stay and discharge destination using hazard functions with unmeasured heterogeneity', *Health Economics* **12**(12), 1021–1034.

Rosen, A. B., Humphries, J., Muhlbaier, L. H., Kiefe, C. I., Kresowik, T. and Peterson, E. D. (1999), 'Effect of clinical factors on length of stay after coronary artery bypass surgery: results of the cooperative cardiovascular project', *American Heart Journal* **138**(1), 69–77.

Rosenberg, M. A., Andrews, R. W. and Lenk, P. J. (1999), 'A hierarchical bayesian model for predicting the rate of nonacceptable in-patient hospital utilization', *Journal of Business & Economic Statistics* **17**(1), 1–8.

Röttger, J., Scheller-Kreinsen, D. and Busse, R. (2012), 'Patient-level hospital costs and length of stay after conventional versus minimally invasive total hip replacement: a propensity-matched analysis', *Value in Health* **15**(8), 999–1004.

Savage, E. and Wright, D. J. (2003), 'Moral hazard and adverse selection in Australian private hospitals: 1989–1990', *Journal of Health Economics* 22(3), 331–359.

Schmid, C. H. and Brown, E. N. (2000), 'Bayesian hierarchical models', *Methods in Enzymology* **321**, 305–330.

Segal, S. P., Preston, N., Kisely, S. and Xiao, J. (2009), 'Conditional release in western Australia: effect on hospital length of stay', *Psychiatric Services* **60**(1), 94–99.

Snijders, T. A. and Bosker, R. J. (1994), 'Modeled variance in two-level models', *Sociological Methods & Research* 22(3), 342–363.

Sørensen, T. H., Olsen, K. R. and Gyrd-Hansen, D. (2009), 'Differences in general practice initiated expenditures across Danish local health authorities-a multilevel analysis', *Health policy* **92**(1), 35–42.

Sundararajan, V., Henderson, T., Perry, C., Muggivan, A., Quan, H. and Ghali, W. A. (2004), 'New ICD-10 version of the charlson comorbidity index predicted in-hospital mortality', *Journal of Clinical Epidemiology* **57**(12), 1288–1294.

Thompson, T. J., Smith, P. J. and Boyle, J. P. (1998), 'Finite mixture models with concomitant information: assessing diagnostic criteria for diabetes', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **47**(3), 393–404.

Weingarten, S., Riedinger, M., Sandhu, M., Bowers, C., Ellrodt, A. G., Nunn, C., Hobson, P. and Greengold, N. (1998), 'Can practice guidelines safely reduce hospital length of stay? results from a multicenter interventional study', *The American journal of medicine* **105**(1), 33–40.

Wright, S., Verouhis, D., Gamble, G., Swedberg, K., Sharpe, N. and Doughty, R. (2003), 'Factors influencing the length of hospital stay of patients with heart failure', *European Journal of Heart Failure* **5**(2), 201–209.

Xu, R. (2003), 'Measuring explained variation in linear mixed effects models', *Statistics in Medicine* **22**(22), 3527–3541.

Yau, K. K., Lee, A. H. and Ng, A. S. (2003), 'Finite mixture regression model with random effects: application to neonatal hospital length of stay', *Computational Statistics & Data Analysis* **41**(3), 359–366.

Zhang, X., Hauck, K. and Zhao, X. (2013), 'Patient safety in hospitals–a bayesian analysis of unobservable hospital and specialty level risk factors', *Health Economics* **22**(9), 1158–1174.

Appendix A

Table A.1: WIES table 2009-2015

WIES	AR-DRG	Inlier Low	Boundary High	Same day weight	One day weight	Multi-day low outlier per diem	Inlier weight	High outlier per diem
WIES 16 (2009-2010)	I03A	4	42	2.0974	2.4860	0.5829	4.8175	0.1590
	I03B	2	22	2.6201	3.0597	0.4395	3.9387	0.1593
WIES 17 (2010-2011)	I03A	4	39	2.0779	2.4642	0.5795	4.7822	0.1699
	I03B	2	21	2.6148	3.0656	0.4508	3.9672	0.1752
WIES 18 (2011-2012)	I03A	3	34	2.2300	2.6784	0.5979	4.4720	0.1694
	I03B	2	19	2.6797	3.0811	0.4014	3.8839	0.1720
WIES 19 (2012-2013)	I03A	3	34	2.2299	2.6783	0.5979	4.4719	0.1694
	I03B	2	19	2.6796	3.0811	0.4014	3.8839	0.1720
WIES 20 (2013-2014)	I03A	3	29	2.2989	2.7163	0.5566	4.3861	0.1764
	I03B	1	16	2.9803	3.6778	0.0000	3.6778	0.1803
WIES 21 (2014-2015)	I03A	3	28	2.1059	2.5025	0.5288	4.0889	0.1832
	I03B	1	15	2.9505	3.7019	0.0000	3.7019	0.2132

Appendix B

Table B.1: Number of episodes in each hospital and year

hosp	1	2	3	4	5	6	7	8	9	10	11	12	Total
h_1	75	76	68	85	98	116	147	166	173	162	152	169	1,487
h_2	82	78	131	127	122	113	242	203	219	217	243	276	2,053
h_3	114	118	115	86	86	105	150	141	153	130	121	97	1,416
h_4	98	76	71	64	73	91	129	113	124	125	138	154	1,256
h_5	64	58	98	123	108	133	185	185	197	237	248	264	1,900
h_6	40	37	38	46	64	57	118	119	112	103	135	123	992
h_7	65	56	61	64	74	82	118	158	146	133	120	96	1,173
h_8	31	18	31	31	35	32	72	82	78	64	62	41	577
h_9	62	101	82	46	47	50	72	72	69	67	53	83	804
h_{10}	110	67	100	74	109	117	159	151	151	125	176	154	1,493
h_{11}	23	36	30	14	17	14	57	65	49	25	47	77	454
h_{12}	52	72	102	108	86	115	364	340	322	300	315	325	2,501
h_{13}	89	94	70	78	66	95	136	136	125	159	162	176	1,386
h_{14}	59	74	51	63	65	60	175	143	152	186	160	198	1,386
h_{15}	70	107	96	96	117	163	209	225	210	229	230	179	1,931
h_{16}	28	53	33	43	31	37	85	82	75	87	113	117	784
h_{17}	58	32	30	34	42	37	100	102	83	85	103	111	817
h_{18}	44	64	64	55	53	48	104	103	98	100	158	128	1,019
h_{19}	97	125	93	81	130	115	151	195	258	217	212	226	1,900
h_{20}	99	98	88	95	99	105	253	216	219	224	282	279	2,057
p_1	56	46	42	23	13	21	51	45	73	64	73	65	572
p_2	153	164	170	168	188	198	602	607	624	724	753	798	5,149
p_3	65	105	112	76	94	121	369	401	373	424	402	424	2,966
p_4	38	31	32	30	27	32	77	66	70	77	84	96	660
p_5	18	18	13	12	24	26	85	66	65	42	52	63	484
p_6	31	35	26	27	41	38	124	115	111	125	131	121	925
p_7	40	88	61	81	44	74	203	210	238	245	278	282	1,844
p_8	29	30	33	34	26	38	196	191	195	207	275	287	1,541
p_9	26	21	36	13	26	16	82	92	90	102	94	105	703
p_{10}	39	45	31	33	50	15	115	139	151	130	101	116	965
p_{11}	16	30	38	30	42	25	78	79	59	62	55	56	570
p_{12}	33	51	41	41	44	51	131	87	124	135	207	277	1,222
p_{13}	132	138	164	152	154	126	670	662	748	709	744	729	5,128
p_{14}	57	64	64	78	86	83	232	240	319	307	329	357	2,216
p_{15}	47	47	50	46	77	72	263	302	385	387	454	582	2,712
p_{16}	36	49	19	26	36	18	136	133	152	168	192	156	1,121
p_{17}	16	21	11	8	11	21	62	58	51	61	87	90	497
p_{18}	9	12	10	15	20	15	64	79	56	89	99	126	594
p_{19}	30	27	21	21	14	14	44	71	76	75	76	73	542
p_{20}	109	119	120	130	98	124	206	233	222	245	265	247	2,118
Total	2,340	2,581	2,546	2,457	2,637	2,813	6,816	6,873	7,195	7,353	7,981	8,323	59,915

Appendix C

The Gibbs sampling procedure is implemented using 10,000 iterations with a burn-in period of 2,000 iterations. The full conditionals are derived as follows.

$$p(\gamma|\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y) \propto \exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\Big(y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\Big)^2\right\}, \tag{11}$$

$$p(\delta_1|\gamma,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y) \propto \exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\Big(y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\Big)^2\right\}, \tag{12}$$

$$p(\delta_2|\gamma,\delta_1,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y) \propto \exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\Big(y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\Big)^2\right\}, \tag{13}$$

$$p(\beta_0|\gamma,\delta_1,\delta_2,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y) \propto \exp\left\{-\frac{1}{\sigma_u^2}\left[\frac{1}{2}\sum_{h=1}^{H}(\beta_h-\beta_0)^2\right]\right\}\exp\left\{-\frac{(\beta_0-\mu_1)^2}{2\sigma_1^2}\right\}, \tag{14}$$

$$p(\alpha_0|\gamma,\delta_1,\delta_2,\beta_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y) \propto \exp\left\{-\frac{1}{\sigma_v^2}\left[\frac{1}{2}\sum_{h=1}^{H}(\alpha_h-\alpha_0)^2\right]\right\}\exp\left\{-\frac{(\alpha_0-\mu_2)^2}{2\sigma_2^2}\right\}, \tag{15}$$

$$p(\sigma_e^2|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_u^2,\sigma_v^2,\beta,\alpha,y) \propto (\sigma_e^2)^{-\frac{N}{2}-a_e-1}\exp\left\{-\frac{1}{\sigma_e^2}\left[\frac{1}{2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\Big(y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\Big)^2+b_e\right]\right\}, \tag{16}$$

$$p(\sigma_u^2|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_v^2,\beta,\alpha,y) \propto (\sigma_u^2)^{-\frac{H}{2}-a_u-1}\exp\left\{-\frac{1}{\sigma_u^2}\left[\frac{1}{2}\sum_{h=1}^{H}(\beta_h-\beta_0)^2+b_u\right]\right\}, \tag{17}$$

$$p(\sigma_v^2|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\beta,\alpha,y) \propto (\sigma_v^2)^{-\frac{H}{2}-a_v-1}\exp\left\{-\frac{1}{\sigma_v^2}\left[\frac{1}{2}\sum_{h=1}^{H}(\alpha_h-\alpha_0)^2+b_v\right]\right\}, \tag{18}$$

$$p(\beta|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\alpha,y) \propto \exp\left\{-\frac{1}{2\sigma_u^2}\sum_{h=1}^{H}(\beta_h-\beta_0)^2\right\}\exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\Big(y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\Big)^2\right\}, \tag{19}$$

$$p(\alpha|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,y) \propto \exp\left\{-\frac{1}{2\sigma_v^2}\sum_{h=1}^{H}(\alpha_h-\alpha_0)^2\right\}\exp\left\{-\frac{1}{2\sigma_e^2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\Big(y_{iht}-X_{iht}\beta_h-Z_{iht}^\top\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\Big)^2\right\}. \tag{20}$$

Therefore we have

$$\gamma|\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y \sim N\left(w\sum_{i=1}^{N_{ht}}\sum_{h=1}^{H}\sum_{t=1}^{T} Z_{iht}\left(y_{iht}-X_{iht}\beta_h-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\right), w\right), \tag{21}$$

$$\delta_1|\gamma,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y \sim N\left(p\sum_{i=1}^{N_{ht}}\sum_{h=1}^{H}\sum_{t=1}^{6} t\left(y_{iht}-X_{iht}\beta_h-Z_{iht}^{\top}\gamma-\alpha_h\right), p\right), \tag{22}$$

$$\delta_2|\gamma,\delta_1,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y \sim N\left(q\sum_{i=1}^{N_{ht}}\sum_{h=1}^{H}\sum_{t=7}^{T} t\left(y_{iht}-X_{iht}\beta_h-Z_{iht}^{\top}\gamma-\alpha_h\right), q\right), \tag{23}$$

$$\beta_0|\gamma,\delta_1,\delta_2,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y \sim N\left(\frac{\sigma_1^2\sum_{h=1}^{H}\beta_h+\sigma_u^2\mu_1}{H\sigma_1^2+\sigma_u^2}, \frac{\sigma_1^2\sigma_u^2}{H\sigma_1^2+\sigma_u^2}\right), \tag{24}$$

$$\alpha_0|\gamma,\delta_1,\delta_2,\beta_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,\alpha,y \sim N\left(\frac{\sigma_2^2\sum_{h=1}^{H}\alpha_h+\sigma_v^2\mu_2}{H\sigma_2^2+\sigma_v^2}, \frac{\sigma_2^2\sigma_v^2}{H\sigma_2^2+\sigma_v^2}\right), \tag{25}$$

$$\sigma_e^2|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_u^2,\sigma_v^2,\beta,\alpha,y \sim IG\left(\frac{N}{2}+a_e, \frac{1}{2}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{i=1}^{N_{ht}}\left(y_{iht}-X_{iht}\beta_h-Z_{iht}^{\top}\gamma-\alpha_h-\delta_1 t-(\delta_2-\delta_1)tI(t\geq 7)\right)^2+b_e\right), \tag{26}$$

$$\sigma_u^2|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_v^2,\beta,\alpha,y \sim IG\left(\frac{H}{2}+a_u, \frac{1}{2}\sum_{h=1}^{H}(\beta_h-\beta_0)^2+b_u\right), \tag{27}$$

$$\sigma_v^2|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\beta,\alpha,y \sim IG\left(\frac{H}{2}+a_v, \frac{1}{2}\sum_{h=1}^{H}(\alpha_h-\alpha_0)^2+b_v\right), \tag{28}$$

$$\beta|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\alpha,y \sim N\left(\mu_\beta,\Sigma_\beta\right), \tag{29}$$

$$\alpha|\gamma,\delta_1,\delta_2,\beta_0,\alpha_0,\sigma_e^2,\sigma_u^2,\sigma_v^2,\beta,y \sim N\left(\mu_\alpha,\Sigma_\alpha\right), \tag{30}$$

where

$$w = \left(\sum_{i=1}^{N_{ht}} \sum_{h=1}^{H} \sum_{t=1}^{T} Z_{iht} Z_{iht}^{\top} \right)^{-1}, \tag{31}$$

$$p = \left(\sum_{i=1}^{N_{ht}} \sum_{h=1}^{H} \sum_{t=1}^{6} t^2 \right)^{-1}, \tag{32}$$

$$q = \left(\sum_{i=1}^{N_{ht}} \sum_{h=1}^{H} \sum_{t=7}^{T} t^2 \right)^{-1}, \tag{33}$$

$$\mu_\beta = \left[\frac{\sigma_u^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} x_{i1t} \left(y_{i1t} - Z_{i1t}^{\top} \gamma - \alpha_1 - \delta_1 t - (\delta_2 - \delta_1) t I(t \geq 7) \right) + \beta_0 \sigma_e^2}{\sigma_u^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} x_{i1t}^2 + \sigma_e^2}, \ldots, \frac{\sigma_u^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} x_{iHt} \left(y_{iHt} - Z_{iHt}^{\top} \gamma - \alpha_H - \delta_1 t - (\delta_2 - \delta_1) t I(t \geq 7) \right) + \beta_0 \sigma_e^2}{\sigma_u^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} x_{iHt}^2 + \sigma_e^2} \right], \tag{34}$$

$$\mu_\alpha = \left[\frac{\sigma_v^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} \left(y_{i1t} - Z_{i1t}^{\top} \gamma - X_{i1t} \beta_1 - \delta_1 t - (\delta_2 - \delta_1) t I(t \geq 7) \right) + \alpha_0 \sigma_e^2}{\sigma_v^2 N_1 + \sigma_e^2}, \ldots, \frac{\sigma_v^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} \left(y_{iHt} - Z_{iHt}^{\top} \gamma - X_{iHt} \beta_H - \delta_1 t - (\delta_2 - \delta_1) t I(t \geq 7) \right) + \alpha_0 \sigma_e^2}{\sigma_v^2 N_H + \sigma_e^2} \right], \tag{35}$$

$$\Sigma_\beta = \mathrm{diag}\left(\frac{\sigma_e^2 \sigma_u^2}{\sigma_u^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} x_{i1t}^2 + \sigma_e^2}, \ldots, \frac{\sigma_e^2 \sigma_u^2}{\sigma_u^2 \sum_{i=1}^{N_{ht}} \sum_{t=1}^{T} x_{iHt}^2 + \sigma_e^2} \right), \tag{36}$$

$$\Sigma_\alpha = \mathrm{diag}\left(\frac{\sigma_e^2 \sigma_v^2}{\sigma_v^2 N_1 + \sigma_e^2}, \ldots, \frac{\sigma_e^2 \sigma_v^2}{\sigma_v^2 N_H + \sigma_e^2} \right). \tag{37}$$

Appendix D

Table D.1: Simulated Inefficiency Factors

Parameter	SIF	Parameter	SIF	Parameter	SIF
β_1	1.7	β_{34}	1.5	α_{27}	58.5
β_2	1.2	β_{35}	0.8	α_{28}	55.9
β_3	1.5	β_{36}	0.8	α_{29}	51.2
β_4	2.4	β_{37}	1.3	α_{30}	57.1
β_5	1.6	β_{38}	1.3	α_{31}	52.1
β_6	2.1	β_{39}	1.0	α_{32}	76.4
β_7	1.4	β_{40}	1.8	α_{33}	86.3
β_8	1.8	α_1	78.4	α_{34}	82.3
β_9	2.2	α_2	82.5	α_{35}	82.1
β_{10}	1.6	α_3	78.6	α_{36}	76.8
β_{11}	1.2	α_4	75.2	α_{37}	36.2
β_{12}	10.4	α_5	82.9	α_{38}	50.8
β_{13}	0.8	α_6	57.2	α_{39}	38.9
β_{14}	1.7	α_7	79.4	α_{40}	82.1
β_{15}	1.3	α_8	48.3	γ_1	44.4
β_{16}	1.8	α_9	53.7	γ_2	2.2
β_{17}	2.3	α_{10}	79.0	γ_3	4.7
β_{18}	1.0	α_{11}	49.3	γ_4	1.8
β_{19}	4.9	α_{12}	83.2	γ_5	2.2
β_{20}	1.8	α_{13}	65.6	γ_6	14.1
β_{21}	2.9	α_{14}	70.7	γ_7	6.0
β_{22}	1.3	α_{15}	67.8	γ_8	1.8
β_{23}	2.0	α_{16}	50.1	γ_9	1.6
β_{24}	2.7	α_{17}	50.0	γ_{10}	1.0
β_{25}	0.9	α_{18}	65.5	γ_{11}	19.5
β_{26}	0.8	α_{19}	78.2	γ_{12}	10.4
β_{27}	1.4	α_{20}	72.3	δ_1	17.4
β_{28}	1.5	α_{21}	49.2	δ_2	23.4
β_{29}	1.5	α_{22}	85.5	σ_e^2	1.6
β_{30}	1.6	α_{23}	84.3	σ_u^2	1.4
β_{31}	1.3	α_{24}	54.0	σ_v^2	3.3
β_{32}	0.8	α_{25}	36.6	β_0	1.7
β_{33}	1.4	α_{26}	44.6	α_0	61.5

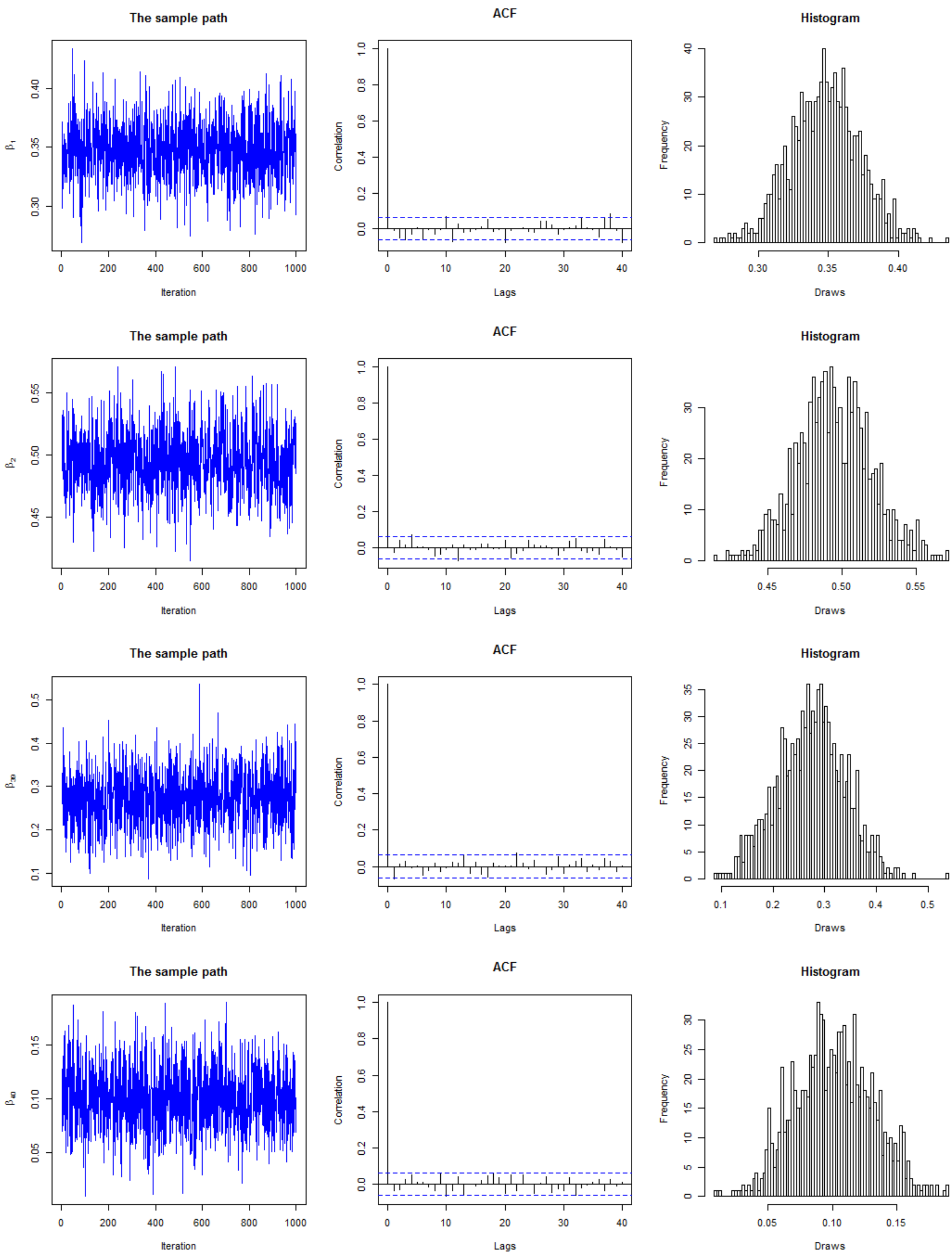


Figure D.1: Sample paths, correlograms and histograms for simulated parameters for patient and hospital characteristics

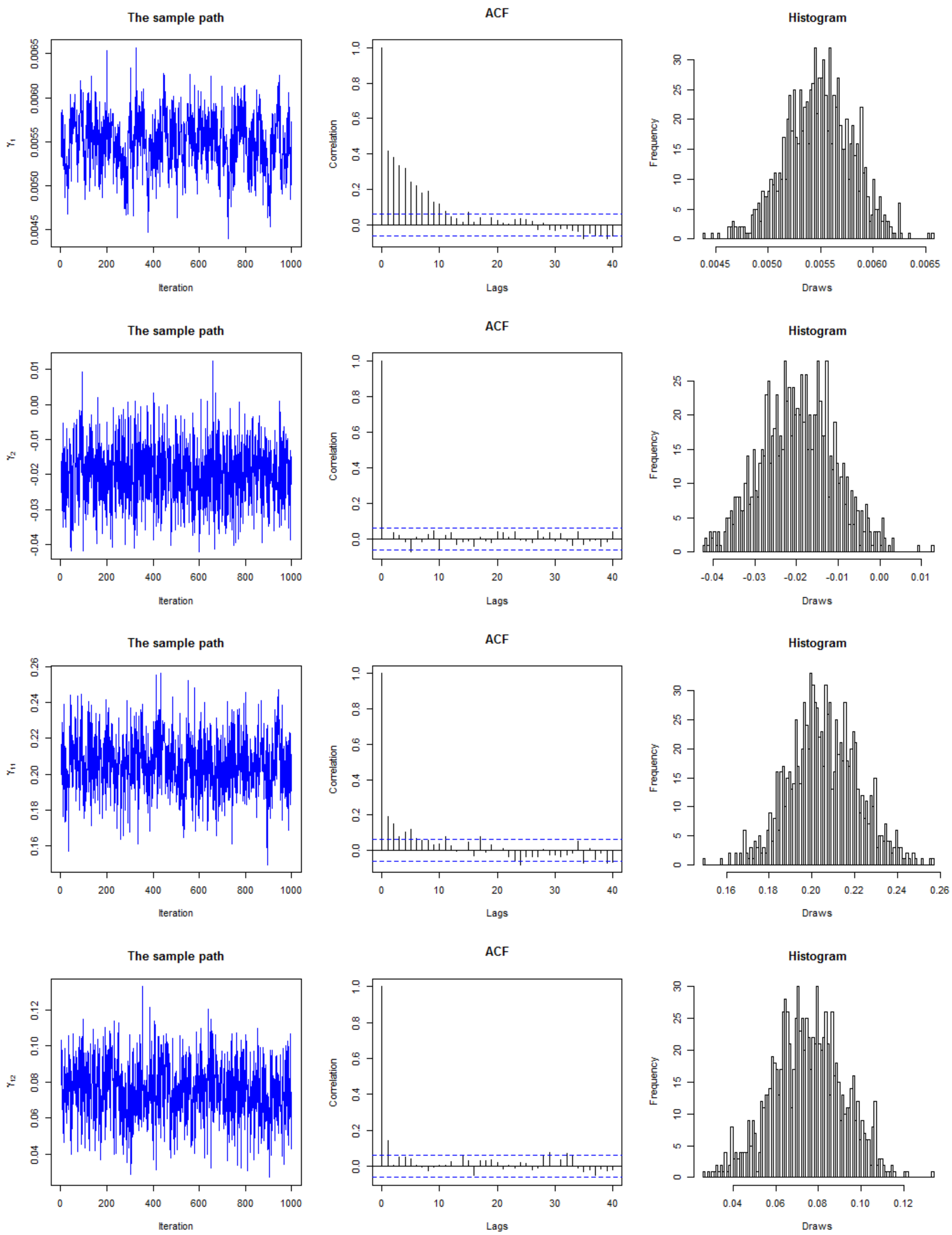


Figure D.2: Sample paths, correlograms and histograms for simulated parameters for patient and hospital characteristics

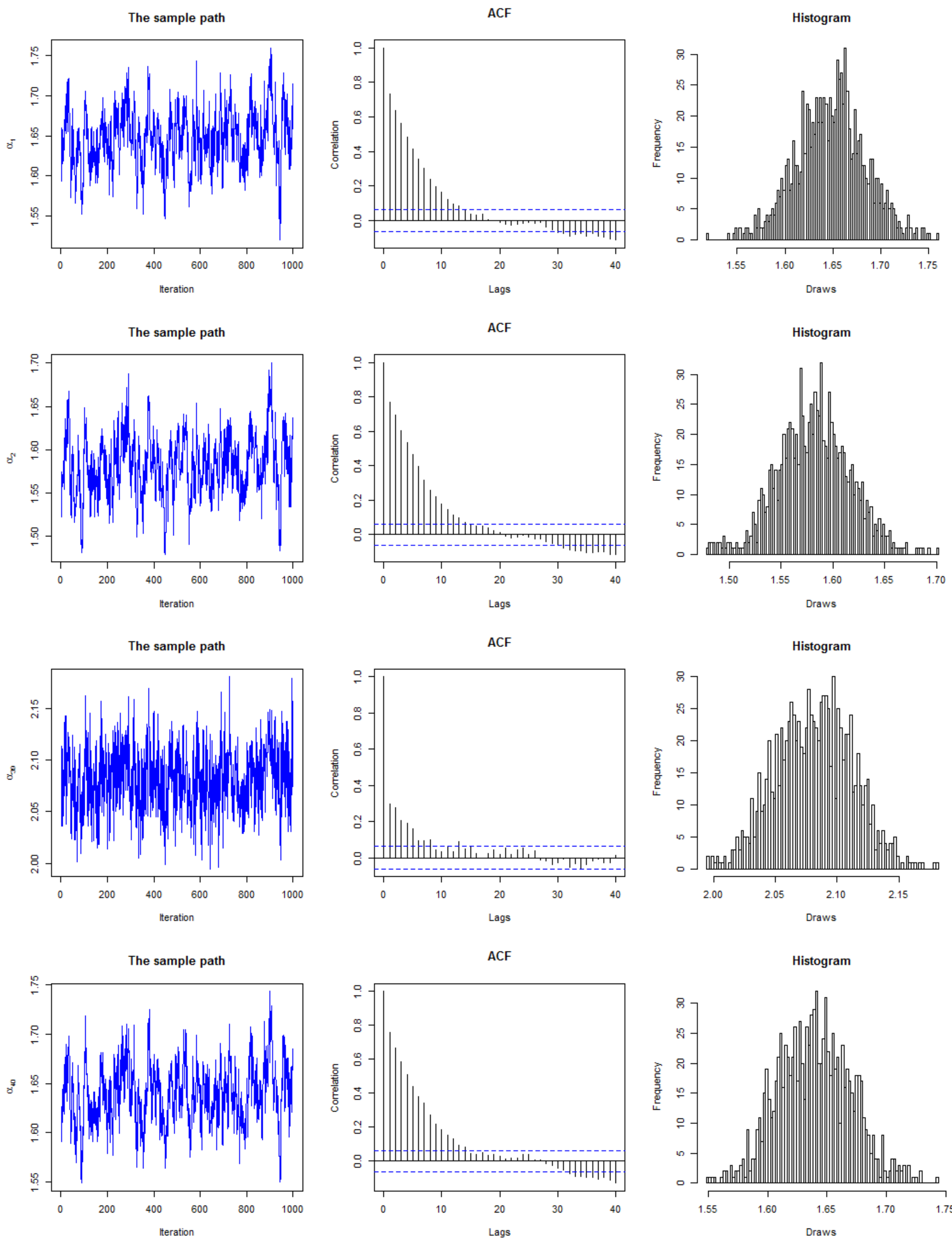


Figure D.3: Sample paths, correlograms and histograms for simulated parameters for public hospital effects

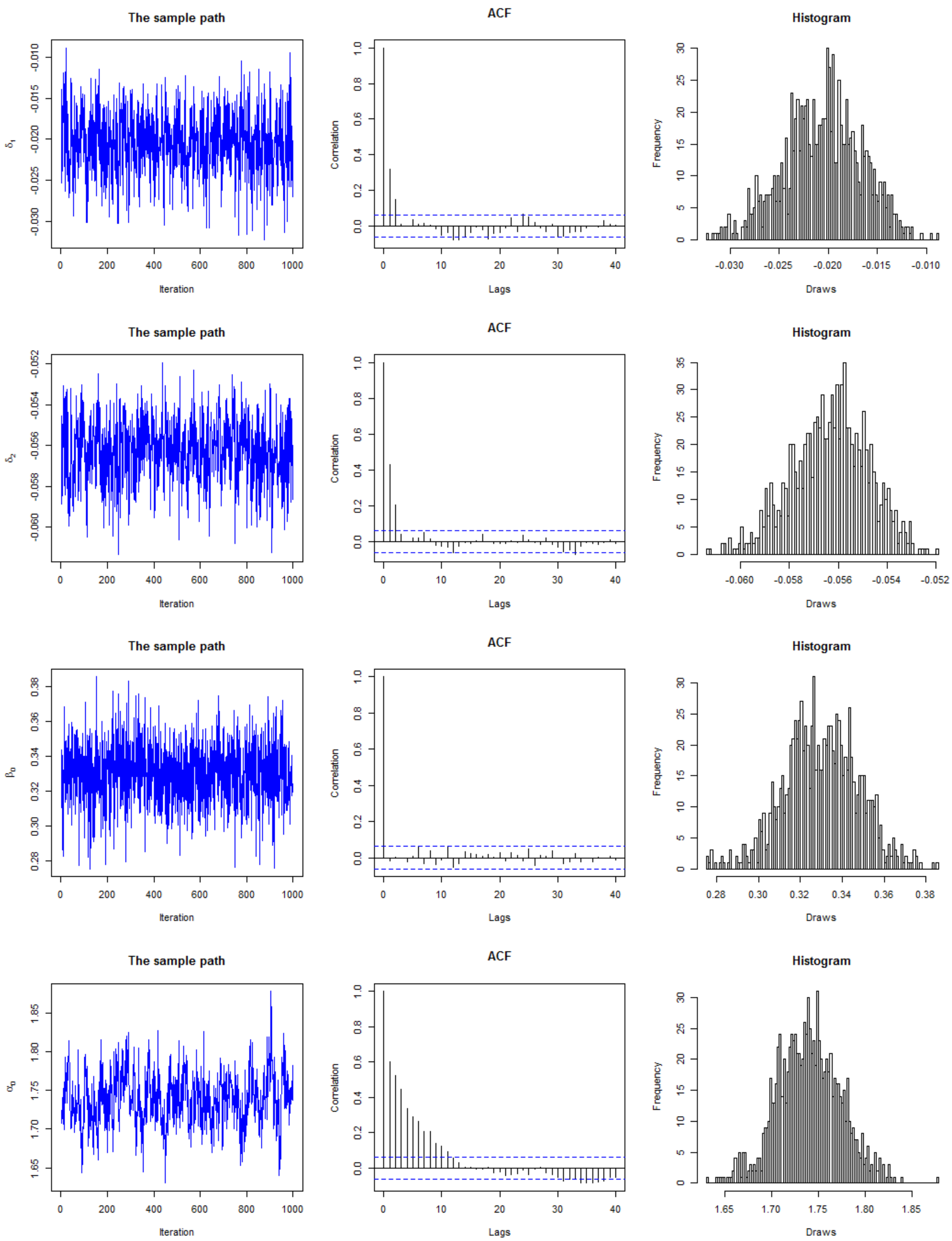


Figure D.4: Sample paths, correlograms and histograms for simulated parameters for public hospital effects

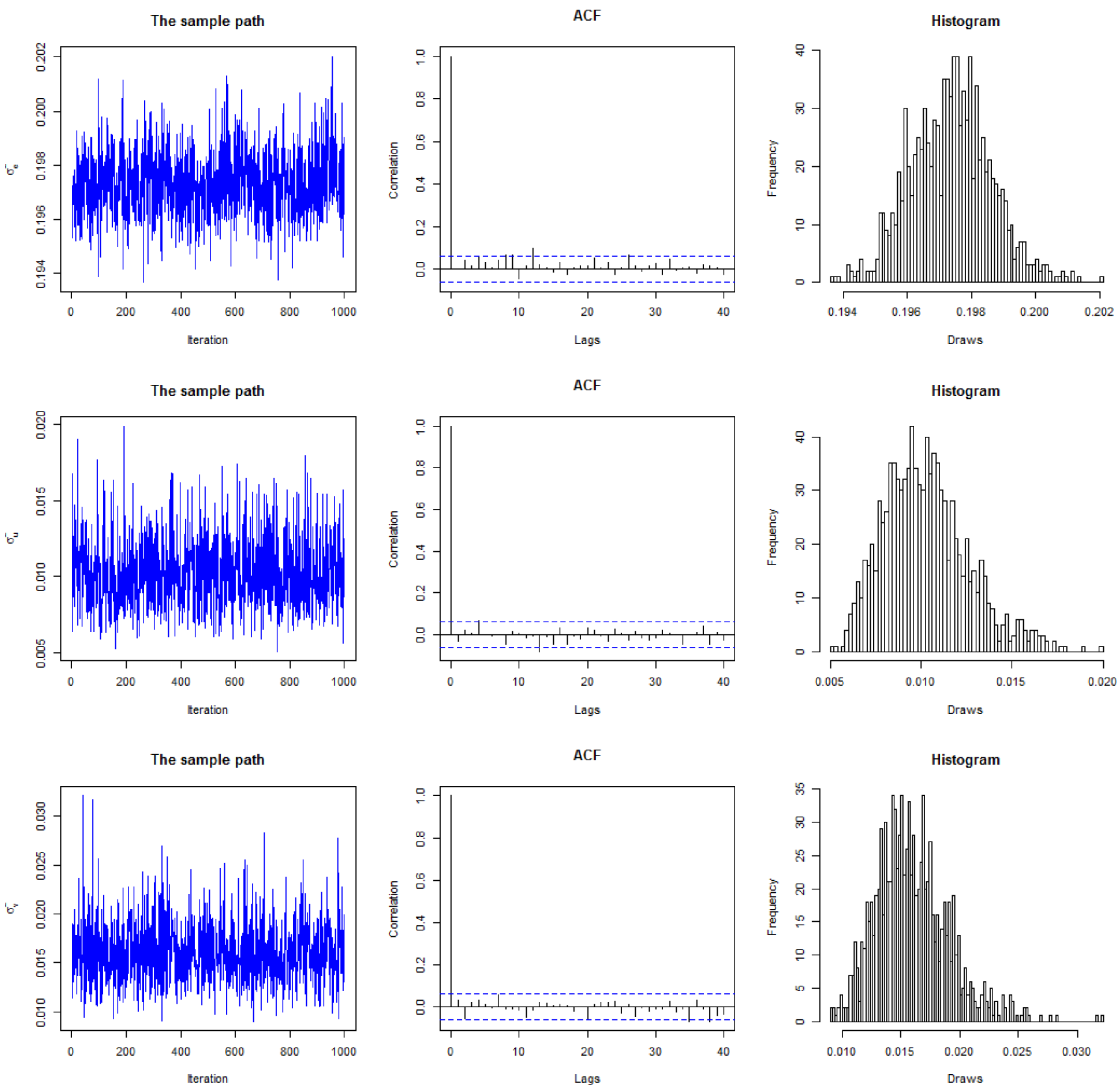


Figure D.5: Sample paths, correlograms and histograms for simulated parameters for public hospital effects