On the Hospital Volume and Outcome Relationship: Does Specialization Matter More Than Volume?

Kris C. L. Lee, Kannan Sethuraman, and Jongsay Yong

Objective. To evaluate the relationship between hospital volume and outcome by focusing on alternative measures of volume that capture specialization and overall throughput of hospitals.

Data Sources/Study Setting. Hospital administrative data from the state of Victoria, Australia; data contain 1,798,474 admitted episodes reported by 135 public and private acute-care hospitals.

Study Design. This study contrasts the volume–outcome relationship using regression models with different measures of volume; two-step and single-step risk-adjustment methods are used.

Data Collection/Extraction Methods. The sample is restricted to ischemic heart disease (IHD) patients (ICD-10 codes: I20–I25) admitted during 2001/02 to 2004/05.

Principal Findings. Overall hospital throughput and degree of specialization display more substantive implications for the volume–outcome relationship than conventional caseload volume measure. Two-step estimation when corrected for heteroscedasticity produces comparable results to single-step methods.

Conclusions. Different measures of volume could lead to vastly different conclusions about the volume–outcome relationship. Hospital specialization and throughput should both be included as measures of volume to capture the notion of size, focus, and possible congestion effects.

Key Words. Hospital volume and outcome, risk adjustment, hospital quality, specialization

This study investigates the association between volume and outcome in the delivery of health services in the context of ischemic heart disease (IHD), the most common type of heart disease and cause of heart attacks. The investigation primarily focuses on the way volume is measured. In studying the volume–outcome relationship, the existing literature tends to focus on specific diseases such as cancer (Gruen et al. 2009) or operative procedures (e.g.,

coronary artery bypass grafting; see Hanchate et al. 2010). In these settings, volume is typically taken as the caseload of the particular disease or operative procedures under study. This study, however, compares the impact of the conventional measure of volume on the outcome with alternate measures focusing on the overall throughput of a hospital as well as the magnitude of its “specialization”; the latter is defined as the proportion of caseload volume for the disease investigated to the overall throughput of the hospital. Thus, a hospital may be low volume but highly specialized if it concentrates its efforts in select areas.

The comparison of different volume measures and their impact on outcome is investigated using different methods of risk adjustment. Specifically, we compare the results under two broad classes of common risk-adjustment methods: two-step versus single-step approaches. The former tends to be popular in health services and epidemiological research, while the latter is more favored in health economics literature.

Over the past three decades, a large body of research in health services research and health economics has investigated the relationship between volume of health services provided by hospitals and patient outcomes (e.g., Luft, Bunker, and Enthoven 1979; Gaynor 2006; Gaynor and Town 2011). This study uses the 30-day mortality after admission as the outcome indicator so as to avoid unfair comparisons due to hospital differences in length of stay (Nicholl, Jacques, and Campbell 2013; Pouw et al. 2013).

It has long been recognized that the case mix differences between hospitals, that is, the differences in patient health conditions at admission, have consequences for the variations in reported performances of the hospitals. Higher risk patients typically generate increased costs even with efficient care providers. In addition, the observed rates of patient outcomes will be systematically influenced by the characteristics of patients, such as age, sex, and presence of comorbidities at admission. It is hence generally agreed that caution must be exercised in making inferences about quality of hospitals and appropriate risk-adjustment approaches accounting for the variations due to the clinical, demographic, and case-mix differences between patients at the point of arrival in hospitals are recommended for fair interinstitutional comparisons (Iezzoni

Address correspondence to Kannan Sethuraman, Ph.D., Melbourne Business School, University of Melbourne, 200 Leicester Street, Carlton, VIC 3053, Australia; e-mail: k.sethuraman@mbs.edu. Kris C. L. Lee, Ph.D., is with the Golden Dragon Centre, City University of Macau, Macau, China, and the Faculty of Business and Economics, University of Melbourne, Melbourne, VIC, Australia. Jongsay Yong, Ph.D., is with the Faculty of Business and Economics, University of Melbourne, Melbourne, VIC, Australia.

2003). Much less, however, is agreed on the appropriate method of risk adjustment (e.g., Silber et al. 2010). In health services research, the risk-adjustment method is typically separate from the intended analysis, which, in this context, is the effect of volume on mortality outcome. The estimation process thus involves two separate steps—a risk-adjustment step followed by a second step dealing with the intended analysis. In contrast, studies in health economics tend to favor a single-step approach, where the risk adjustment and intended analysis are combined through a single-step estimation (e.g., Picone, Trogon, and Jollis 2005; Tsai et al. 2006), although there have also been studies using the two-step approaches (e.g., Street et al. 2012 in a study of hospital utilization).

The literature on the association between hospital volume and outcome mostly supports the thesis that high-volume hospitals have significantly better health outcomes for their patients compared to their low-volume counterparts, after adjusting for patient risk factors (Hannan, 1999; Birkmeyer et al. 2002; Halm, Lee, and Chassin 2002; Shahian and Normand 2003; Ho 2002; Gaynor, Seider, and Vogt 2005). For instance, in the case of pancreatic resection, Birkmeyer et al. (2002) showed that risk-adjusted mortality rates at very low-volume hospitals (those averaging less than one pancreatic resection per year) were 12.5 percent higher than the rates at very high-volume hospitals (16.3 percent vs. 3.8 percent). Purchasers and consumer groups, such as the Leapfrog group, have supported the wider public dissemination of data on hospital volume for specific procedures and cite the support for positive association between hospital volume and outcome as a rationale for advocating selective referral to high-volume hospitals and surgeons (Birkmeyer, Finlayson, and Birkmeyer 2001; Birkmeyer and Dudley 2014).

Two key explanations for this positive relationship have been offered: the “practice makes perfect” and “selective-referral” effects (Luft, Hunt, and Maerki 1987). The former hypothesis treats volume as an exogenous attribute and links the better quality outcomes at the larger hospitals to the learning-by-doing effect. This effect essentially captures the learning that accrues to the physicians in hospitals that treat a larger number of patients with similar conditions in comparison to those in lower volume hospitals. Selective-referral effect, on the other hand, treats volume as an endogenous attribute in that hospitals known for their higher quality care are likely to attract more referrals and thus accrue a larger volume of patients. A number of strategies have been suggested to tackle the contaminating effect of this reverse causality issue (e.g., Gowrisankaran, Ho, and Town 2006; Tsai et al. 2006; Huesch 2009).

This study does not aim at addressing the causality issue; rather, it suggests an alternative view on volume—that volume may not only capture the learning-by-doing effect but also the size effect, the scope and scale economies accruing through the comprehensiveness of services provided at a hospital, and the degree of specialization. A hospital that provides a comprehensive set of clinical services may provide better outcomes for their patients due to both “static economies of scope,” that is, benefits of related diversification, and complementarities that accrue due to new insight and knowledge about related conditions (Clark and Huckman 2011). For example, a diabetic patient with heart disease would be better cared for in a large, comprehensive care provider that can not only offer cardiac care but also provide advice on diabetes management through staff dietitians on its roster. The new knowledge and insight gained through the interactions with other clinical specialists could potentially help cardiovascular specialists arrive at better interpretations, helping to improve quality of outcomes for their patients. Thus, the size of the hospital acting as a surrogate for capturing the scope economies arising out of the comprehensiveness of its services could be an important contributor to the quality of care it provides. Likewise, the extent to which a hospital specializes in serving patients with a particular illness may be an important determinant of outcome; for example, specialty hospitals, such as women’s or children’s hospitals, could be better equipped and staffed to handle complex cases in their respective specialty areas than general hospitals. This reasoning builds on the focused-factory approach of Skinner (1974), who advocates that specialization improves performance through a firm’s ability to focus on fewer areas with greater effectiveness; Herzlinger (1997) and Greenwald et al. (2006) have explored this notion in the context of the health care sector.

Another important dimension of volume that has not received much attention in the literature is that hospital throughput may act as a surrogate for capturing the implication of congestion faced by large, public hospitals in countries such as Australia, where hospital care is provided via a tax-funded universal health care system. This consideration is particularly relevant for public hospitals in countries where funding constraints often result in congestion, long waiting times, and crowded facilities; see Siciliani and Hurst (2005) for a review of the waiting time phenomenon in the case of elective surgery.

In summary, volume could play several roles that are not necessarily complementary to one another. This study thus proposes using overall throughput as a measure of hospital size and using the proportion of IHD episodes to all admission episodes as a measure of its degree of specialization.

In the estimations below, these measures are found to compare favorably to the conventional measure of IHD caseload volume.

DATA

The data source for this study is hospital administrative data from the State of Victoria, Australia. The database, known as the Victorian Admitted Episode Dataset, contains detailed information on admitted-patient episodes reported by all public and private acute hospitals in the state. The data include demographic, clinical, and administrative details for all admitted episodes of care occurring in Victorian acute hospitals. The data have been linked to the death registry via a statistical linking process developed by the Victorian Department of Human Services.

The full sample consisted of 1,798,474 admission episodes generated by all IHD patients admitted to 303 hospitals during a 7-year period from 1998/99 to 2004/05. The full sample of admission episodes was used for developing the risk-adjustment model in the first step of our two-step estimation methods. For the second step, the estimation was restricted to 135 hospitals; the restricted sample excluded hospitals that were identified as day clinics¹ and small hospitals with an annual throughput below 1,000 admissions or with fewer than four IHD episodes a year. Day clinics were excluded because they specialize in minor surgical procedures (e.g., cataract removal), for which few if any deaths were observed. Small hospitals and hospitals dealing with fewer than four IHD episodes a year were also excluded for the same reason; the number of deaths usually were relatively few in these hospitals, and a slight change in the number of deaths typically exaggerates the impact on the calculated standardized mortality rates of these hospitals, resulting in considerable noise to the estimation if they had been included. In addition to the above restrictions, we also further restricted the sample to cover only the admission episodes during the 4 years from 2001/02 to 2004/05 so as to minimize the impact of clinical and technological advancements on clinical outcomes. To make the estimation comparable between the two-step and single-step methods, we also applied these restrictions to the latter. Thus, for the single-step estimation, the sample only contained 1,035,323 admission episodes occurring in 135 hospitals.

The admission episodes were identified with the use of appropriate medical diagnosis codes from the International Classification of Diseases, 10th Revision (ICD-10 codes I20–I25). We used the outcome measure, mortality

within 30 days after admission, to avoid “discharge bias” (Nicholl, Jacques, and Campbell 2013; Pouw et al. 2013). For risk adjustment, we used both patient- and admission-specific characteristics. The former included patient characteristics such as their age, gender, marital status, and private health insurance coverage status, whereas the latter included clinical variables designed to capture the severity and/or complexity of the admission episodes. Among covariates of a clinical nature was the Charlson Comorbidity index, which measures the complexity of an admission episode and has been shown to be a good predictor of mortality (Charlson et al. 1987). We computed the Charlson Comorbidity index based on the procedure outlined in Sundararajan et al. (2004).

At the hospital-level, besides the volume measures, additional hospital-level covariates included teaching hospital status, proportion of private patients treated, and the number of competing public and private hospitals in the hospital’s catchment area, the last being defined using the procedure proposed by Melnick and Zwanziger (1988); see also Palangkaraya and Yong (2013).

Table 1 summarizes the dependent variable and covariates used in the estimation. The sample consisted of relatively older patients, with an average age of 69.8 years. The 30-day mortality rate for the whole sample was 1.8 percent. There were a total of 135 hospitals, including 53 private hospitals in the sample. Hospitals in the sample had on average 12,687 admissions a year, but their range varied from 1,027 to 93,122 admissions. These hospitals treated on average 405 IHD admissions a year, with a range that varied from 5 to 3,079 IHD admissions. The proportion of IHD admissions to total admissions ranged from 0.02 to 14.4 percent with an average of 2.8 percent. The correlation coefficient between total volume and proportion of IHD admissions was found to be 0.12, and thus unlikely to cause any collinearity problems.

The graphs in Figure 1 depict the data in scatterplots for two primary determinants—caseload volume and specialization—in relation to risk-adjusted mortality rates; also included in each plot are the locally weighted scatterplot smoothing estimator (LOWESS) curve and least-squares line. Two different risk-adjusted mortality rates were used, namely the rate ratio (of observed-to-expected mortality) and the rate difference (between the observed and expected mortality). Consistent with previous findings (e.g., Spiegelhalter et al. 2012), the plots of caseload volume against risk-adjusted mortality show greater variability for low-volume hospitals in comparison to their high-volume counterparts, while generally the observed mortality for

Table 1: Summary Statistics on Dependent Variable and Covariates

	<i>Mean</i>	<i>Standard Deviation</i>	<i>Median</i>	<i>IQR</i>
Admission characteristics				
30-day mortality (from admission date)	0.018	0.134	0.000	0.000
Charlson Comorbidity Index	1.228	1.677	1.000	2.000
Number of diagnoses in last admission	3.415	2.736	2.000	2.000
Had heart bypass operation before	0.064	0.245	0.000	0.000
Had angioplasty before	0.094	0.292	0.000	0.000
Had chemotherapy before	0.052	0.221	0.000	0.000
Had dialysis before	0.339	0.473	0.000	1.000
Admitted via emergency department	0.261	0.439	0.000	1.000
Same-day separation	0.57	0.495	1.000	1.000
Transferred from other hospital	0.055	0.229	0.000	0.000
Had private insurance	0.291	0.454	0.000	1.000
Male	0.611	0.487	1.000	1.000
Married	0.556	0.497	1.000	1.000
Australian born	0.603	0.489	1.000	1.000
Age	69.752	11.953	72.000	16.000
Number of admissions	1,798,474			
Hospital characteristics				
Throughput	1.269	1.548	0.665	1.415
IHD volume	0.041	0.067	0.009	0.036
IHD specialization	0.028	0.025	0.025	0.023
Teaching hospital status	0.164	0.371	0.000	0.000
Proportion of private patients	0.369	0.379	0.143	0.779
Number of competing public hospitals	9.065	7.912	7.000	10.000
Number of competing private hospitals	21.576	22.976	15.000	34.000
Number of hospitals	135			

Throughput = total number of annual admission episodes in a hospital. IHD volume = total number of IHD caseload volume in a hospital. IHD specialization = IHD volume as proportion of total throughput.

the high-volume hospitals tend to match their expected mortality. Although there were no obvious trends visually observable in the association between volume/specialization and risk-adjusted mortality, both the LOWESS curves and least-squares lines suggest slight downward-sloping trends, indicating lower than expected mortality rates being associated with increased volume or specialization.

EMPIRICAL SPECIFICATION

To investigate the sensitivity of volume–outcome relationship with different measures of volume, we conducted two separate sets of analyses for each

Figure 1: Scatterplots of Caseload Volume and Specialization versus Mortality (Respective LOWESS Curves and Least-Squares Lines Are Shown)

Notes: (1) Extreme values are excluded in the above diagrams: rate differences above 0.1 and rate ratios above 3 are excluded. (2) Estimated least-squares lines in subfigures (a)–(d) are as follows: (a) Rate Diff = $0.0025 - 0.0007 \times \text{Volume}$; t -stat for slope is -1.13 ($p = .26$). (b) Rate Diff = $0.0036 - 0.07 \times \text{Specialization}$; t -stat for slope is -1.91 ($p = .06$). (c) Rate Ratio = $1.081 - 0.018 \times \text{Volume}$; t -stat for slope is -1.01 ($p = .31$). (d) Rate Ratio = $1.142 - 2.98 \times \text{Specialization}$; t -stat for slope is -2.77 ($p = .006$).

empirical specification described below. In the first set of analysis, the conventional volume measure, IHD caseload volume, was used, while in the second set of analysis, hospital throughput and IHD specialization were used to characterize volume. We evaluated the volume–outcome relationship over several common specifications in the literature to establish the robustness of the results.

All empirical specifications below began with the following two equations as the building blocks. Let i , h , and t index patients, hospitals, and time periods, respectively. Let y_{iht} denote the mortality outcome (0 = survived, 1 = dead) for the i th patient treated at the h th hospital during the t th time period.

$$\Pr(y_{iht} = 1 | X_{iht}) = g(X_{iht}\beta + \eta_{ht}) \quad (1)$$

$$\eta_{ht} = Z_{ht}\gamma + v_{ht} \quad (2)$$

In the above model, equation (1) states that the probability of mortality outcome can be expressed as a function of patient-specific admission characteristics, X_{iht} and an unobserved hospital-level attribute term, η_{ht} . The function $g(\cdot)$ is typically specified as logistic, although it may also be expressed as a linear probability model (e.g., Tsai et al. 2006). Equation (2) characterizes the hospital-level attribute, η_{ht} as a linear function of some observable hospital characteristics, Z_{ht} and an *i.i.d.* error term v_{ht} . Included in the vector Z are measures of hospital volume, the variable of primary interest in this study.

In a two-step risk-estimation process, equations (1) and (2) were estimated separately. Equation (1), with admission episode as the unit of observations, constitutes the risk-adjustment step. Using the estimates from step 1, an estimate of η_{ht} is determined and used in equation (2), which has hospital as the unit of observations.

In the empirical implementation below, two variants of equation (1) were considered. In the first variant, a logistic specification (without the η_{ht} term) was estimated. From the estimated equation, we obtained the predicted probability of mortality for the i th patient treated at the h th hospital in the t th time period, $\hat{p}_{iht} = g(X_{iht}\hat{\beta})$, for each admission episode. The expected number of mortalities at the h th hospital in the t th time period, $E_{ht} = \sum_{i \in h} \hat{p}_{iht}$ was then computed. The risk-adjusted mortality rate of a hospital was expressed in one of the two ways: (i) as a difference between the hospital's observed number of deaths and its expected number of deaths, $O_{ht} - E_{ht}$, or (ii) as a ratio of the hospital's observed to its expected number of deaths, $\frac{O_{ht}}{E_{ht}}$. In step 2, the adjusted mortality rate was then used as the dependent variable as per equation (2).

In the second variant of implementing equation (1), a linear probability model with hospital fixed effects interacting with year dummies was specified, and the estimated hospital and year effects $\hat{\eta}_{ht}$ were extracted and then used as the dependent variable in equation (2). We did not estimate a logistic regression with hospital dummy variables because of the incidental parameters problem and convergence issues due to hospitals with few or no deaths.

In two-step estimation, because the dependent variable for equation (2) is an estimated quantity, the loss of efficiency in OLS estimation due to heteroscedastic errors is a concern. Following Street et al. (2012), we applied the feasible generalized least-squares estimator (FGLS) of Lewis and Linzer

(2005) in all the step 2 estimations to allow for heteroscedastic errors. As an alternative, we also attempted the Huber–White sandwich estimator; results were similar and are available upon request.

In single-step estimation, equation (2) is substituted into equation (1) to obtain a two-level estimating equation of the form:

$$Pr(y_{iht} = 1|X_{iht}) = g(X_{iht}\beta + Z_{ht}\gamma + v_{ht}) \quad (3)$$

Equation (3) was implemented in two ways. First, it was implemented as a random-intercept logistic regression where the hospital-specific error term, v_{ht} , is modeled as normally distributed with zero mean and variance σ_v . Second, as unobserved individual patient characteristics could be correlated with the error terms, we also introduced Mundlak-adjustment terms (Mundlak 1978) into equation (3).

RESULTS

Table 2 shows the results obtained from the two-step risk-adjustment methods where the standard errors were estimated using Lewis–Linzer FGLS estimator (Lewis and Linzer 2005); we also estimated the standard errors using the Huber–White sandwich estimator method (Huber 1967; White 1980)² and the results obtained were similar. In Table 2, the top half of the table displays estimates from using IHD caseload volume as the volume measure, while the bottom half presents the estimates from using hospital throughput and IHD specialization as the volume measures. The first-step estimation results can be found in Table 3.

With volume measured using IHD caseload volume, we found no statistically significant relationship between risk-adjusted mortality and volume. However, when volume was measured using the alternate measures, namely, hospital throughput and IHD specialization, a statistically significant relationship (at 5 percent or higher significance level) was found between IHD specialization and risk-adjusted mortality. In other words, an increase in specialization of a hospital was significantly associated with a fall in risk-adjusted mortality. This is consistent with our earlier observation on the scatterplots. The relationship was found for all the cases whether we used rate difference, rate ratio, or fixed effects estimates as the dependent variables. We also observed a weak relationship between hospitals' teaching status and risk-adjusted mortality indicating lower levels of risk-adjusted mortality rates at teaching hospitals, although the relationship was not statistically significant.

Table 2: Two-Step Methods, Step 2 Estimates, Standard Errors Estimated by Lewis–Linzer FGLS Estimator

	<i>Diff</i> ($O_j - E_j$)		<i>Ratio</i> (O_j/E_j)		<i>Linear FE</i>	
	<i>Coeff.</i>	<i>SE</i>	<i>Coeff.</i>	<i>SE</i>	<i>Coeff.</i>	<i>SE</i>
Volume measured by IHD caseload volume						
IHD volume	−0.0055	0.0150	0.1244	0.4494	−0.0114	0.0102
Teaching hospital status	−0.0031	0.0031	−0.0821 [†]	0.0942	−0.0018	0.0029
Proportion of private patients	−0.0064**	0.0031	−0.0723	0.0927	−0.0051*	0.0026
Number of competing public hospitals	−0.0001	0.0003	−0.0124 [†]	0.0075	0.00001	0.0002
Number of competing private hospitals	0.0001	0.0001	0.0021	0.0028	0.0001*	0.0001
Volume measured by throughput and specialization						
Throughput	−0.0002	0.0007	0.0213	0.0211	−0.0004	0.0007
IHD specialization	−0.0730*	0.0333	−3.0461***	0.9922	−0.0711**	0.0232
Teaching hospital status	−0.0027	0.0033	−0.1143	0.0995	−0.0025	0.0032
Proportion of private patients	−0.0059 [†]	0.0031	−0.0508	0.0915	−0.0054*	0.0026
Number of competing public hospitals	−0.0001	0.0002	−0.0114	0.0074	0.00004	0.0002
Number of competing private hospitals	0.0001	0.0001	0.0019	0.0028	0.0001	0.0001
Number of observations	493		493		493	
Number of hospitals	135		135		135	

Note: Dependent variables are constructed using 30-day mortality since admission; standard errors were estimated using Lewis–Linzer FGLS estimator (Lewis and Linzer 2005).

Significance levels: [†]10%, *5%, **1%, ***0.1%.

There was also a weak negative association between proportion of private patients and risk-adjusted mortality, implying that the higher the proportion of private patients, the lower the levels of risk-adjusted mortality.

Table 4 summarizes the estimation results using the single-step approach. For ease of comparison with the previous results, marginal effect and elasticity estimates were presented. Also presented were model fit statistics that included log-likelihood values, AIC and BIC statistics, and pseudo R-squared statistics; the last was computed as the proportion of the increase in log likelihood of the current model against the intercept-only model.

Given that the random effects logistic model is nested in the Mundlak-adjusted model, we also conducted a likelihood ratio test. The test produced chi-squared statistics of 96.54 and 107.76; the former was for the specification including only IHD caseload volume and the latter was obtained when both throughput and specialization were used as volume measures. The test results

Table 3: Two-Step Methods, Step 1 Risk-Adjustment Models

<i>Variable</i>	<i>Logit</i>		<i>Linear FE</i>	
	<i>Coeff.</i>	<i>SE</i>	<i>Coeff.</i>	<i>SE</i>
Charlson Comorbidity Index	0.3274***	0.0023	0.0130***	0.0001
Number of diagnoses in last September	0.0530***	0.0017	0.0014***	0.00004
Had heart bypass operation before	−0.4680***	0.0352	−0.0044***	0.0004
Had angioplasty before	−0.4686***	0.0293	−0.0061***	0.0004
Had chemotherapy before	−0.0936**	0.0282	−0.0206***	0.0005
Had dialysis before	−0.4153***	0.0277	−0.0012***	0.0003
Admitted via emergency department	1.3494***	0.0148	0.0307***	0.0003
Same-day separation	−1.2438***	0.0218	−0.0123***	0.0003
Transferred from other hospital	0.2326***	0.0204	0.0027***	0.0005
Had private insurance	0.0947***	0.0141	0.0031***	0.0003
Male	0.1342***	0.0127	0.0023***	0.0002
Married	−0.0399**	0.0129	−0.0006**	0.0002
Australian born	0.0825***	0.0129	0.0004	0.0002
Age 46–55	0.4123***	0.1018	0.0026***	0.0006
Age 56–65	0.9368***	0.0946	0.0039***	0.0006
Age 66–75	1.4085***	0.0927	0.0070***	0.0006
Age 76–85	2.0314***	0.0923	0.0191***	0.0006
Age above 85	2.7442***	0.0926	0.0558***	0.0007
SEIFA disadvantage index	0.0047*	0.0021	0.0001	0.00004
ARIA Remoteness index	−0.0140*	0.0059	0.0005**	0.0002
Intercept	−6.9318***	0.0956	−0.0192***	0.0044
Number of observations	1,798,474		1,671,007	
Log likelihood	−124,273		—	
Pseudo R^2 /Adj. R^2	0.2477		0.0631	

Significance levels: *5%, **1%, ***0.1%.

clearly rejected the null; thus, the Mundlak-adjusted model was the preferred specification in this context.

Similar to the results in Table 2, a statistically significant negative relation between IHD specialization and risk-adjusted mortality was observed in the results in Table 4. This negative relationship was particularly strong for the Mundlak-adjusted model. However, we found no statistically significant relationship when volume was measured using IHD caseload volume.

Importantly, results under the Mundlak-adjusted model showed a positive and moderately significant relationship between hospital throughput and risk-adjusted mortality. The elasticity estimate for hospital throughput suggested that a 1 percent increase in throughput would lead to a 0.17 percent increase in risk-adjusted mortality, other things being equal. In contrast, the corresponding elasticity estimate for IHD specialization suggested that a 1 percent increase in IHD specialization would be associated with a 0.013

Table 4: Single-Step Logistic Estimation, Selected Estimates

	<i>RE Logistic</i>			<i>Mundlak-Adjusted RE Logistic</i>		
	<i>Marg. Eff.</i>	<i>SE</i>	<i>Elasticity</i>	<i>Marg. Eff.</i>	<i>SE</i>	<i>Elasticity</i>
Volume measured by IHD caseload volume						
IHD volume	−0.0018	0.0022	−0.0444	−0.0011	0.0023	−0.0264
Teaching hospital status	0.0001	0.0006	0.0161	0.0007	0.0007	0.0848
Proportion of private patients	0.0004	0.0007	0.0219	0.0002	0.0011	0.0122
Number of competing public hospitals	−0.00003	0.00002	−0.0860	−0.00004	0.00002	−0.0908
Number of competing private hospitals	0.000003	0.00001	0.0185	−0.00001	0.00001	−0.0380
Log likelihood		−74,406.0			−74,357.8	
AIC		148,866.0			148,797.5	
BIC		149,186.0			149,283.4	
Pseudo R^2		0.340			0.341	
Likelihood ratio test, H_0 : All Mundlak adj. terms = 0	Chi-squared statistics = 96.54 (p = .0000), reject H_0					
Volume measured by throughput and specialization						
Throughput	0.0000	0.0001	0.0283	0.0003 [†]	0.0001	0.1713*
IHD specialization	−0.0095 [†]	0.0053	−0.0062 [†]	−0.0206**	0.0063	−0.0131**
Teaching hospital status	−0.0001	0.0006	−0.0071	0.0004	0.0007	0.0408
Proportion private patients	0.0004	0.0007	0.0209	−0.0001	0.0012	−0.0074
Number of competing public hospitals	−0.00003	0.00002	−0.0722	−0.00003	0.00003	−0.0752
Number of competing private hospitals	−0.000001	0.00001	−0.0037	−0.00001	0.00001	−0.0795
Log likelihood		−74,404.7			−74,350.8	
AIC		148,865.4			148,785.7	
BIC		149,197.2			149,283.4	
Pseudo R^2		0.340			0.341	
Likelihood ratio test, H_0 : All Mundlak adj. terms = 0	Chi-squared statistics = 107.76 (p = .0000), reject H_0					
Number of observations		1,035,323			1,035,323	
Number of hospitals		135			135	

Notes: Pseudo R -squared is computed as the proportion of the increase in log likelihood of the current model against the intercept-only model.
Significance levels: [†]10%, *5%, **1%.

percent decline in risk-adjusted mortality. The effects of hospital throughput and IHD specialization appeared to work in opposite directions, with the effect of the former dominating the effect of the latter in elasticity terms.

DISCUSSION

This study evaluates the hospital volume–outcome relationship pertaining to IHD patients admitted to hospitals in Victoria, Australia, during a 4-year period covering the financial years 2001/02 to 2004/05. The evaluation focuses on whether the volume–outcome relationship is affected by the use of different measures of volume. To check the robustness of our results, we compared estimation results of commonly deployed two-step methods with single-step random effects logistic estimation.

The main finding is that alternative measures of volume alter the results in a substantive way. The use of two volume measures, namely hospital throughput and IHD specialization, produced consistent results that suggest higher degrees of specialization were associated with lower risk-adjusted mortality rates. The conventional single-volume measure relying on IHD caseload volume produced no statistically significant results in all our estimations. Further, single-step logistic regressions provide an interesting insight that risk-adjusted mortality may be positively associated with hospital throughput while being negatively associated with IHD specialization. The increase in risk-adjusted mortality with increase in hospital throughput may be indicative of the funding constraints faced by large public hospitals, resulting in long waits and poor quality of care.

As an extension, we also examined the interaction effects between volume and specialization by including an interaction term in all models studied. In all cases, the coefficient estimates of the interaction term were observed to be positive and statistically significant, suggesting the existence of a moderating effect of volume on the favorable effect of specialization on risk-adjusted mortality.³ As it is difficult to disentangle the interaction effect without further investigation, we hypothesize that the congestion effect caused by high volume is possibly negating the positive effect of specialization on mortality, especially in the short term. As this topic is beyond the scope of this paper, we leave further investigation of the interaction effect for future studies.

This study also found that two-step estimation with heteroscedasticity corrected through Lewis–Linzer FGLS estimator could produce comparable results to single-step methods as the efficiency loss of the former is minimized.

In investigating the volume–outcome relationship, based on the results obtained, this study advocates the inclusion of overall hospital throughput as a volume measure, alongside the conventional caseload volume measure, the

latter possibly expressed as a proportion to overall throughput. While single-step estimation methods appeared to produce more efficient estimates than two-step methods, the latter have advantages when the focus is on a select group of hospitals (e.g., large public hospitals). Single-step estimation when constrained to include just the subset of hospitals would not make full use of all information for risk adjustment. Although in principle one could have circumvented this problem by running a multilevel estimation with the use of a dummy variable marking the relevant subset of observations, in practice, this strategy may not work since one often has to exclude hospitals with few or no mortality counts to achieve convergence in estimation. The two-step estimation procedure, however, allows the use of the complete set of admission episodes from all hospitals during the initial risk-adjustment step, followed by the estimation on the subset of hospitals of interest in the second step.

It is worth pointing out that this study only investigates the association between volume and outcome. No attempt has been made to untangle the direction of causation. An understanding of the association is a useful first step not only in understanding the impact of volume in the form of throughput and specialization but also for GPs and health care professionals to decide on hospitals to which patients with particular needs are to be referred.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: Funding from the Australian Research Council Linkage Grant LP0455325 and National Health and Medical Research Council Partnership Grant 567217 is gratefully acknowledged. We thank the Victorian Department of Health (formerly the Victorian Department of Human Services) for providing the data. We also thank two anonymous reviewers for their helpful comments and suggestions.

Disclosures: None.

Disclaimers: None.

NOTES

1. Day clinics are not identified in the data as such. This study defines a day clinic as a hospital with more than 80 percent of its annual admissions as same-day admissions.
2. We also computed robust standard errors by allowing for clustering by hospitals. Unfortunately, the use of two-step methods caused large efficiency losses, leading to

high standard errors and hence few statistically significant estimates; the results are available from the authors upon request.

3. Results are available upon request.

REFERENCES

- Birkmeyer, J. D., and R. A. Dudley. 2014. "The Leapfrog Group: Evidence-Based Hospital Referral Fact Sheet" [accessed on January 17, 2014]. Available at http://www.leapfroggroup.org/media/file/Leapfrog-Evidence-based_Hospital_Referral_Fact_Sheet.pdf
- Birkmeyer, J. D., E. V. Finlayson, and C. M. Birkmeyer. 2001. "Volume Standards for High-Risk Surgical Procedures: Potential Benefits of the Leapfrog Initiative." *Surgery* 130 (3): 415–22.
- Birkmeyer, J. D., A. E. Siewers, S. R. Finlayson, T. A. Stukel, F. Lee Lucas, I. Batista, H. G. Welch, and D. E. Wennberg. 2002. "Hospital Volume and Surgical Mortality in the United States." *New England Journal of Medicine* 346 (15): 1128–37.
- Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKenzie. 1987. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation." *Journal of Chronic Diseases* 40: 373–83.
- Clark, J. R., and R. Huckman. 2011. "*Broadening Focus: Spillovers, Complementarities, and Specialization in the Hospital Industry*." Working Paper 16937. Cambridge, MA: National Bureau of Economic Research.
- Gaynor, M. 2006. *What Do We Know about Competition and Quality in Health Care Markets?* NBER Working Paper 12301. Cambridge, MA: NBER.
- Gaynor, M., H. Seider, and W. B. Vogt. 2005. "The Volume-Outcome Effect, Scale Economies and Learning-By-Doing." *American Economic Review* 95 (2): 243–7.
- Gaynor, M., and R. Town. 2011. *Competition in Health Care Markets*. Working Paper 17208. Cambridge, MA: National Bureau of Economic Research.
- Gowrisankaran, G., V. Ho, and R. J. Town. 2006. *Casualty, Learning and Forgetting in Surgery*. Unpublished Transcript. Tuscon, AZ: University of Arizona.
- Greenwald, L., J. Cromwell, W. Adamache, S. Bernard, E. Drozd, E. Root, and K. Devers. 2006. "Specialty versus Community Hospitals: Referrals, Quality and Community Benefits." *Health Affairs* 25 (1): 106–18.
- Gruen, R. L., V. Pitt, S. Green, A. Parkhill, D. Campbell, and D. Jolley. 2009. "The Effect of Provider Case Volume on Cancer Mortality: Systematic Review and Meta-Analysis." *A Cancer Journal for Clinicians* 59 (3): 192–211.
- Halm, E. A., C. Lee, and M. R. Chassin. 2002. "Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature." *Annals of Internal Medicine* 137 (6): 511–20.
- Hanchate, A. D., T. A. Stukel, J. D. Birkmeyer, and A. S. Ash. 2010. "Surgery Volume, Quality of Care and Operative Mortality in Coronary Artery Bypass Graft Surgery: A Re-Examination Using Fixed-Effects Regression." *Health Services Outcomes Research Method* 10: 16–32.

- Hannan, E. L. 2011. "The Relation between Volume and Outcome in Health Care." *New England Journal of Medicine* 340(21): 1677–9.
- Herzlinger, R. E. 1997. *Market-Driven Health Care*. Reading, MA: Addison-Wesley.
- Ho, V. 2002. "Learning and the Evolution of Medical Technologies: The Diffusion of Coronary Angioplasty." *Journal of Health Economics* 21: 873–85.
- Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions." In *Vol. 1 of Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, edited by L. M. Le Cam and J. Neyman, pp. 221–33. Berkeley, CA: University of California Press.
- Huesch, M. D. 2009. "Learning by Doing, Scale Effects, or Neither? Cardiac Surgeons after Residency." *Health Services Research* 44 (6): 1960–81.
- Iezzoni, L. I. 2003. *Risk Adjustment for Measuring Health Care Outcomes*. Chicago, IL: Health Administration Press.
- Lewis, J. B., and D. A. Linzer. 2005. "Estimating Regression Models in Which the Dependent Variable is Based on Estimates." *Political Analysis* 13: 345–64.
- Luft, H. S., J. P. Bunker, and A. C. Enthoven. 1979. "Should Operations Be Regionalized? The Empirical Relation between Surgical Volume and Mortality." *New England Journal of Medicine* 301: 1364–9.
- Luft, H. S., S. S. Hunt, and S. B. Maerki. 1987. "The Volume-Outcome Relationship: Practice-Makes-Perfect or Selective-Referral Patterns?" *Health Services Research* 22 (2): 157–82.
- Melnick, G. A., and J. Zwanziger. 1988. "Hospital Behavior under Competition and Cost-Containment Policies: The California Experience, 1980-1985." *Journal of the American Medical Association* 260: 2669–75.
- Mundlak, Y. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46: 69–85.
- Nicholl, J., R. Jacques, and M. J. Campbell. 2013. "Mortality Indicators Used to Rank Hospital Performance." *British Medical Journal* 347: f5952.
- Palangkarya, A., and J. Yong. 2013. "Effects of Competition on Hospital Quality: An Examination Using Hospital Administrative Data." *European Journal of Health Economics* 14 (3): 415–29.
- Picone, G., J. G. Trogon, and J. Jollis. 2005. *Hospital Volume and Quality of Care: Selective-Referral or Practice-Makes-Perfect?* Working Paper. Tampa, FL: University of South Florida.
- Pouw, M., L. Peelen, K. Moons, C. Kalkman, and H. Linsma. 2013. "Including Post-Discharge Mortality in Calculation of Hospital Standardized Mortality Ratios: Retrospective Analysis of Hospital Episode Statistics." *British Medical Journal* 347: f5913.
- Shahian, D. M., and S. L. Normand. 2003. "The Volume-Outcome Relationship: From Luft to Leapfrog." *Annals of Thoracic Surgery* 75: 1048–58.
- Siciliani, L., and J. Hurst. 2005. "Tackling Excessive Waiting Times for Elective Surgery: A Comparative Analysis of Policies in 12 OECD Countries." *Health Policy* 72: 201–15.
- Silber, J. H., P. R. Rosenbaum, T. J. Brachet, R. N. Ross, L. J. Bressler, O. Even-Shoshan, S. A. Lorch, and K. G. Volpp. 2010. "The Hospital Compare Mortality

- Model and the Volume-Outcome Relationship.” *Health Services Research* 45 (5): 1148–67.
- Skinner, W. 1974. “The Focused Factory.” *Harvard Business Review* 52 (3): 113–21.
- Spiegelhalter, D., C. Sherlaw-Johnson, M. Bardsley, I. Blunt, C. Wood, and O. Grigg. 2012. “Statistical Methods for Healthcare Regulation: Rating, Screening and Surveillance.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 175: 1–47.
- Street, A., C. Kobel, T. Renaud, and J. Thuilliez. 2012. “How Well Do Diagnosis-Related Groups Explain Variations in Costs or Length of Stay among Patients and Across Hospitals? Methods for Analyzing Routine Patient Data.” *Health Economics* 21 (Suppl 2): 6–18.
- Sundararajan, V., T. Henderson, C. Perry, A. Muggivan, H. Quan, and W. A. Ghali. 2004. “New ICD-10 Version of the Charlson Comorbidity Index Predicted In-Hospital Mortality.” *Journal of Clinical Epidemiology* 57: 1288–94.
- Tsai, A. C., M. Votruba, J. F. P. Bridges, and R. D. Cebul. 2006. “Overcoming Bias in Estimating the Volume-Outcome Relationship.” *Health Services Research* 41 (1): 252–64.
- White, H. 1980. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica* 48: 817–38.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.