

Hentschker, Corinna; Mennicken, Roman

**Conference Paper**

## Selective-referral and unobserved patient heterogeneity - Bias in the volume-outcome relationship

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Health II, No. B10-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

Suggested Citation: Hentschker, Corinna; Mennicken, Roman (2014) : Selective-referral and unobserved patient heterogeneity - Bias in the volume-outcome relationship, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Health II, No. B10-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<http://hdl.handle.net/10419/100299>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Selective-referral and unobserved patient heterogeneity - Bias in the volume-outcome relationship

Corinna Hentschker & Roman Mennicken\*

August 2013

PRELIMINARY – do not cite without permission of the authors

## **Abstract**

This paper examines the causal effect of volume on outcome on the example of patients with a hip fracture. We use an instrumental variable approach and consider both the practice-makes-perfect and selective-referral hypothesis as well as unobserved patient heterogeneity. Our results indicate that unobserved severity drives the results in the volume-outcome relationship for hip fracture patients and with this the practice-makes-perfect hypothesis has an even stronger effect on hospital quality than expected so far.

KEYWORDS: volume; hospital quality; mortality; instrumental variables

JEL: I11, I12, I18

---

\*Both RWI.

# 1 Introduction

The volume-outcome relationship is a frequently analyzed topic in the literature. In 1979, Luft et al. showed in their seminal paper that in ten out of twelve conditions a correlation between volume and outcome exists (Luft et al., 1979). Numerous other studies followed in the years after. The volume-outcome relationship can be explained by two hypotheses with reversed causality directions: The practice-makes-perfect hypothesis states that higher case volume leads to a better quality because learning affects routines, and economies of scale apply (Luft et al., 1987; Seider et al., 2004). Hospitals which treat more patients with a specific condition reduce their mistakes, optimize processes, and can afford better technical equipment. Hence, volume is the leading cause for good practice. In contrast, the selective-referral hypothesis assumes that good quality hospitals have a higher case volume. This is the result of the good reputation of the hospital: Referring physicians know which hospitals are of good quality and refer patients to a specific hospital. Another reason for this hypothesis could be that patients inform themselves via quality reports and choose the hospital e.g. with the lowest mortality rate. Based on these arguments quality is the leading cause of a high case volume. Both hypotheses are possible and if simultaneously valid, OLS estimation leads to biased and inconsistent results, i.e. the effect of volume is overestimated in a regression of volume on outcome.

Nevertheless most of the previous literature has only focused on the correlation in the volume-outcome relationship and reviews conclude that in general a correlation between volume and outcome exists, but the effect size depends on the condition considered (Gandjour et al., 2003; Halm et al., 2002). Only a few studies tried to determine the causal relationship by using instrumental variable regression or simultaneous equation models. Common instruments are the number of hospital beds (Farley and Ozminkowski, 1992; Luft et al., 1987; Norton et al., 1998) and geographical factors (Barker et al., 2011; Seider et al., 2004; Tay, 1999). Hamilton and Hamilton (1997) use a duration model with hospital fixed effects to exclude selective-referral; however, reverse causality can still drive their results. Barker et al. (2011) and Seider et al. (2004) conclude that the practice-makes-perfect hypothesis is the leading cause for the volume-outcome relationship. In contrast, Tsai et al. (2006) and Hamilton and Hamilton (1997) find evidence for selective-referral as the primary reason in the volume-outcome relationship. Farley and Ozminkowski (1992) and Luft et al. (1987)

provide mixed results, while Norton et al. (1998) and Gaynor et al. (2005)<sup>1</sup> even come to the conclusion that volume is exogenous.

It seems possible that, besides reverse causality, the volume-outcome relationship might be biased by an omitted variable bias due to unobserved patient heterogeneity. Patients may choose hospitals based on quality or even on their initial health status (Tay, 1999). It is obvious that patient characteristics are usually unequally distributed among hospitals, i.e. there are e.g. university hospitals treating sicker patients in terms of age and comorbidities. Problems arise because patient characteristics are correlated with the outcome (Iezzoni, 2003). Therefore, patient characteristics are essential for an appropriate risk adjustment. Most studies analyzing the volume-outcome relationship use administrative data (Halm et al., 2002). Even though these data sets can have very detailed information, clinical parameters like laboratory values, functional status or symptoms and detailed socioeconomic characteristics are missing. Hence, in simple regression models, there might be an unexplained part which can be correlated with the outcome variable and the case volume. For example, if (unobserved) sicker patients are treated more often in high volume hospitals, this would yield to a decline in the measured quality, because those patients have a higher risk to die independent from the quality of the hospital. The coefficient of volume is in this case affected by case volume, which should cause a lower mortality rate, and unobserved illness characteristics which increases mortality.

The issue of unobserved patient heterogeneity in volume-outcome studies is less frequently discussed. Tay (1999) models the effect of volume on outcome for patients with acute myocardial infarction (AMI) with a hospital choice model estimated using a logit model with random coefficients, i.e. they control for the selection of patients with different initial health statuses. Tay (1999) do not find a significant effect of case volume, if the omitted variable bias is ignored. However, when controlling for unobserved heterogeneity, volume becomes significant. Tay (1999) concludes that more severely ill patients choose higher volume hospitals. Hamilton and Hamilton (1997) and Hockenberry et al. (2010) use similar duration models which also allow accounting for unobserved patient severity. They use a generalized estimating equation and model unobserved heterogeneity with a step function. Hockenberry et al. (2010) find that the marginal effects of volume increase after accounting for unobserved heterogeneity.

---

<sup>1</sup>Gaynor et al. (2005) is the corresponding published article of the working paper of Seider et al. (2004).

ity. Further literature which does not directly analyze the volume-outcome relationship also concludes that unobserved patient heterogeneity matters (Becker, 2007; Gowrisankaran and Town, 1999; Kessler and McClellan, 2000). Hence, not only reverse causation is relevant in the volume-outcome relationship but also unobserved patient heterogeneity.

The relevance for assessing the causal effect of volume on outcome arises from the high political impact: The volume-outcome effect is the foundation for minimum volume standards. For example, in Germany minimum volume standards were introduced for five conditions in 2004. Consequently, hospitals which do not achieve a certain number of cases within a specific condition are not allowed to treat patients with this condition anymore.<sup>2</sup> This regulation was introduced following international evidence for the volume-outcome relationship, but most studies only found evidence for a correlation between volume and outcome. However, from a health policy point of view, the causality direction in volume-outcome matters as minimum volume standards implicitly ground on the practice-makes-perfect hypothesis: One of the main concerns against minimum volume standards is that they could endanger access to hospital services, i.e. minimum-volume standards impose a tradeoff between potential gains in quality of care and losses in access to care. If the practice-makes-perfect hypothesis holds, minimum-volume standards will most likely improve overall outcomes, because hospitals treating more cases will improve their quality. However, this implies longer travel times for patients due to the reduced number of hospitals providing treatment. If higher volume does not lead to better quality, i.e. the selective-referral hypothesis applies, minimum-volume standards would be unfavorable as access to care would deteriorate by driving low-volume providers out of the market (Seider et al., 2004). To ensure that the minimum volume regulation really has an effect on quality and not only a specializing effect, the causality direction has to be determined.

We examine the causal effect of volume on outcome for hip fracture patients following Henschker and Mennicken (2012). As outlined, the volume effect might be biased if reverse causality and/or unobserved patient heterogeneity exist. To exclude these biases we instrument case volume with the number of potential patients and the number of further hospitals in the regional area around each hospital. With an instrumental variable approach it is not possible to separate the effect of selective-referral and unobserved patient heterogeneity

---

<sup>2</sup>In 2011 the social court stopped the minimum-volume standard for knee-replacements due to substantial doubts on a causal relationship between quantity and quality.

(Hockenberry et al., 2010) but because they work in different directions, it is possible to determine the stronger influence. So far, there is no study available which determines the causal effect of volume on outcome with German data. Furthermore, there is no standard way on which level the volume-outcome relationship should be analyzed. Three of the seven studies considering causality in volume-outcome use the hospital level as the unit of analysis (Barker et al., 2011; Farley and Ozminkowski, 1992; Luft et al., 1987), while the remaining four use the patient level (Hamilton and Hamilton, 1997; Norton et al., 1998; Seider et al., 2004; Tsai et al., 2006). There is no consensus in the literature, which level is more appropriate. Hence, we analyze the volume-outcome relationship on both levels and compare results.

The remainder of this paper is organized as follows: Section 2 presents the data and first descriptive summaries. Section 3 explains the applied methods; section 4 shows the estimation results, followed by robustness checks and discussion in section 5 and 6. Section 7 summarizes the main findings and concludes.

## 2 Data

We use administrative data of all German hospitals for the year 2007. The data<sup>3</sup> contains the total in-patient population except psychiatric cases in Germany. It includes detailed information on patient level like age, gender, main and secondary diagnosis, procedure codes, admission and discharge reasons, and the zip code of residence of each patient. Furthermore, we have hospital characteristics available e.g. ownership type, teaching status, bed capacity, and the full address of each hospital. We geo-coded the addresses of hospitals and the centroids of all German ZIP codes<sup>4</sup>, so we are able to calculate the distance for each patient to the hospital and distances between hospitals using the Stata command `traveltime` (Ozimek and Miles, 2011).

In this study we concentrate on patients with hip fracture. We use the diagnosis and procedure codes based on the definition of the Federal Office for Quality Assurance (Bundesgeschäftsstelle Qualitätssicherung, 2008). We only include patients with a main diagnosis

---

<sup>3</sup>The data was used as part of the further development of the DRG-system.

<sup>4</sup>We assume that all patients with a specific ZIP code area live at the geographic centroid and patient ZIP codes were based on the home address. Given that the median size of a ZIP code in Germany is about 27 square kilometers, this assumption seems reasonable.

of HIP and a matching procedure code. We exclude patients who are younger than 20 years ( $n = 133$ ) because those patients might need a special treatment compared to older patients. Furthermore, we drop patients who have discharge reason “transfer” ( $n = 9,210$ ). For those patients we are not able to determine the outcome of the treatment. Finally, we exclude patients in hospitals treating less than ten cases in the year 2007 ( $n = 338$ ). We assume that those hospitals do not usually treat hip fracture patients and hence do not belong to the regular provision of hip fracture treatments. Our final sample consists of 89,203 patients treated in 1,155 hospitals.

We use in-hospital mortality<sup>5</sup> as outcome measure. Mortality is the most frequently used outcome measure in volume-outcome studies for two reasons: Mortality has the advantage that it is robust against coding differences between hospitals. This is because every hospital records its own data and coding differences might apply between hospitals but mortality is a clear defined outcome and hence coding differences are not possible. Furthermore, for hip fracture patients mortality is an approved indicator by the Agency for Healthcare Research and Quality (AHRQ) which can be used to determine quality differences between hospitals (Agency for Healthcare Research and Quality, 2007).

Our main explanatory variable is the case volume of each hospital. It varies from 1 to 387 treated hip patients per hospital and year. We assume that patients who are treated in hospitals with a higher case volume have a lower probability of dying in that hospital. In our model we use the logarithm of volume (Farley and Ozminkowski, 1992; Hamilton and Hamilton, 1997).

The outcome of a hospital’s treatment depends not only on the case volume but also on patient risk factors. We use age, gender, admission reason (scheduled, emergency, transfer), and the Charlson comorbidity index (CCI) as control variables. We expect that patients with increasing age, and with more comorbidities have a higher risk to die independent of the quality of the hospital. The CCI is a frequently used measure for the number and severity of comorbidities of a patient (Charlson et al., 1987). Another well established method for risk-adjustment are the Elixhauser comorbidities (Elixhauser et al., 1998). We present results using both methods in separate analyses. We are using diagnosis codes developed by Quan et al. (2005) who mapped the original codes from ICD-9 system to the ICD-10 system used

---

<sup>5</sup>Unfortunately, we are not able to track patients after discharge with the data available. Hence, we cannot consider out-of-hospital mortality.

in Germany. Furthermore, we add dummy variables for admission during winter time<sup>6</sup> as well as for admissions over weekend and public holidays. The first variable captures possible seasonal patterns during winter time. The latter variable captures weekend and public holiday effects due to lower staffing levels in comparison to weekdays (Bell and Redelmeier, 2001). We further differentiate patients with a femoral neck fracture and a pertrochanteric fracture. As another covariate we use a binary variable which indicates whether a transfer between departments during the hospital stay has taken place.

It has been shown that, besides the case volume, other hospital characteristics like e.g. ownership (Milcent, 2005), and teaching status (Ayanian and Weissman, 2002) can influence the quality of a hospital. Hence we include indicator variables for the ownership type, teaching status, university hospital, existence of an intensive care unit (ICU), and bed size of the hospital. It is *ex ante* not clear how mortality might be affected. On the one hand it is possible that these hospital characteristics have a structural effect and lower the probability of death. On the other hand it might be they are correlated with unobserved patient characteristics. If there is a correlation with the case volume it is possible that hospital characteristics increase the risk of mortality.

Table 1 shows descriptive statistics of the patient and hospital characteristics. In total 89,203 patients with hip fracture were treated in 1,155 hospitals. Most of the patients were female (75%) and on average 80 years old. 65% of the patients had one or more Charlson comorbidities. 28% of the patients were admitted on a weekend and 33% during winter time. Around 50% of the patients were treated in a public hospital, 56% were treated in a teaching hospital, and only 4% were treated in a university hospital.

As we have the total in-patient population for Germany we can determine the number of further hospitals and the number of patients in the area around each hospital. These variables serve as our instruments (see section 3) and are similar to Seider et al. (2004). The authors assumed that the hospital case volume depends on the number of patients which live in the regional area of the hospital and on the number of further hospitals in the regional area. Consequently, the case volume of a hospital should increase if more patients with a specific condition live near the hospital and decrease the more hospitals treat this condition in the regional area.

---

<sup>6</sup>This variable is 1, if the admission was in the months November, December, January, or February and 0 otherwise.



Table 1: Descriptive statistics of HIP

	<b>Mean</b>	<b>S.D.</b>
Death	0.063	0.243
ln case volume	4.520	0.566
Femoral neck fracture	0.537	0.499
Male	0.251	0.434
Age	79.564	11.290
Admission reason		
Scheduled admission	0.218	0.413
Emergency	0.762	0.426
Transfer	0.020	0.141
Transfer between departments	0.271	0.444
Winter	0.331	0.471
Weekend or holiday admission	0.284	0.451
Charlson comorbidity index		
0	0.357	0.479
1-2	0.433	0.496
3-4	0.153	0.360
$\geq 5$	0.057	0.232
Ownership		
Public	0.494	0.500
Private non-profit	0.376	0.484
Private for-profit	0.131	0.337
University hospital	0.038	0.190
Teaching hospital	0.556	0.497
Beds		
200 beds	0.197	0.398
201-499 beds	0.478	0.500
$\geq 500$ beds	0.325	0.468
ICU	0.448	0.497
Number of patients	89,203	
Number of hospitals	1,155	

Seider et al. (2004) use the straight-line distances and choose different radii for the potential patients and the number of further hospitals. We have calculated real travel times by car in minutes. With travel times by car we are able to account for differences in the infrastructure. We use the number of potential patients who live 0 to 10 minutes, 10 to 20 minutes and 20 to 30 minutes away from the hospital and the number of further hospitals up to 15 minutes and between 15 and 30 minutes.<sup>7</sup> Furthermore, population and hospital density show substantial variation throughout Germany. For this reason we add indicators for the settlement structures (SST) to our instrument (BBSR - Bundesinstitut für Bau Stadt- und Raumforschung, 2012). These indicators differentiate county types by population density and settlement structure, i.e. they range from agglomerated regions with cities without county membership and with more than 100,000 inhabitants (SST 1) over urbanized regions with more than 150 inhabitants per square kilometers (SST 6) to rural regions with less than 100 inhabitants per square kilometers (SST 9). A detailed definition for all SST is given in the Appendix (see Table A1).

Table 2 shows descriptive statistics of the instruments for whole Germany and differentiated by SST. Taking the perspective of a hospital, on average two further hospitals are within a 15 minutes' drive. There are notable differences between the SST—the number of further hospitals reaches from six in SST 1, over 2 hospitals in SST 2 and SST 5, to 0 hospitals in SST 3, SST 4, SST 6-9. The number of further hospitals correlates positive with the number of potential patients. On average, a hospital can find nearly 100 hip fracture patients within ten minutes in Germany. In SST 1 there are already 210 patients within this radius compared to less than 30 patients in SST 9.

---

<sup>7</sup>These thresholds are somewhat arbitrary. We estimated models with different thresholds but the models stayed basically the same (see section 5).

Table 2: Number of potential patients and number of further hospitals by SST

SST	Number of patients	Number of hospitals	Number of potential patients between ...			Number of further hospitals between ...	
			0 to 10 min	10 to 20 min	20 to 30 min	0 to 15 min	15 to 30 min
			SST 1	22,200	247	210	767
SST 2	13,052	163	88	429	964	2	16
SST 3	6,139	78	43	150	461	0	6
SST 4	2,506	42	33	86	302	0	5
SST 5	7,568	76	121	205	259	2	4
SST 6	16,735	227	48	114	256	1	4
SST 7	8,424	143	29	50	154	0	3
SST 8	8,372	107	42	63	135	0	2
SST 9	4,207	72	27	35	84	0	1
Overall	89,203	1,155	96	318	580	2	10

### 3 Methods

We specify our dependent variable  $y_{ih}$  as a binary variable which indicates whether patient  $i$  has died in hospital  $h$ . We estimate the following equation via OLS<sup>8</sup>:

$$y_{ih} = \alpha_0 + \ln(vol)_h \beta_{A1} + x'_{ih} \beta_{A2} + k'_h \beta_{A3} + \varepsilon_{ih} \quad (1)$$

where  $\ln(vol)_h$  is the logarithm of the case volume,  $x_{ih}$  are patient characteristics,  $k_h$  are hospital characteristics, and  $\varepsilon_{ih}$  is a random error term. Equation (1) specifies the volume-outcome relationship as practice-makes-perfect hypotheses. The parameter  $\beta_{A1}$  is biased by reversed causality (selective-referral) and unobserved patient heterogeneity. Hence we have to eliminate reverse causality and unobserved heterogeneity and find an instrument which is strongly correlated with the case volume but not with the quality of a hospital and thus with the error term. Accordingly, we only use the variation of case volume which is uncorrelated with the error term.

In general, patients choose hospitals which are closer to their residence (Seider et al., 2004). This implies that the case volume of a hospital depends on the number of potential patients  $p_h$  and the number of further hospitals  $h_h$  in the regional area around each hospital which we are using as an instrument (see section 2). Conversely, the quality of a hospital should not have an influence on the potential patients or the number of further hospitals.

<sup>8</sup>Usually models with a dependent binary variable are estimated with a logit or a probit model. We estimate all our presented models also in a probit specification (see Appendix Table A2). The interpretation stays the same.

Patients' residences can be considered as exogenous to the hospital quality, because it is unlikely that patients choose residency on the basis of quality of care in a nearby hospital. Furthermore, other hospital locations can be viewed as exogenous, because a hospital closure or opening is a rare event.<sup>9</sup>

Furthermore, the instruments must not be correlated with unobserved patient heterogeneity. We follow Gowrisankaran and Town (1999) by assuming that unobserved severity of illnesses is uniformly distributed in the population, i.e. the number of potential patients should not be influenced by unobserved heterogeneity.

We use a two stage least square (2SLS) estimation to eliminate selective-referral and unobserved severity in equation (1) and focus only on practice-makes-perfect. We specify the first stage equation (2), where the logarithm of the case volume is regressed on all covariates of equation (1) and the instruments  $p_h$  and  $h_h$  to get  $\widehat{\ln(vol)}_{ih}$ . In the second stage equation (3) the fitted values of  $\ln(vol)_{ih}$  from equation (2) are used to model the causal effect of volume on outcome.

$$\ln(vol')_{ih} = \alpha_0 + \mathbf{x}'_{ih}\boldsymbol{\pi}_{A1} + \mathbf{k}'_h\boldsymbol{\pi}_{A2} + \mathbf{p}'_h\boldsymbol{\gamma}_{A1} + \mathbf{h}'_h\boldsymbol{\gamma}_{A2} + \mathbf{s}'_h\boldsymbol{\gamma}_{A3} + v_{ih} \quad (2)$$

$$y_{ih} = \alpha_0 + \widehat{\ln(vol')}_{ih}\beta_{B1} + \mathbf{x}'_{ih}\boldsymbol{\beta}_{B2} + \mathbf{k}'_h\boldsymbol{\beta}_{B3} + v_{ih} \quad (3)$$

The second stage equation (3) differs from the structural model (1) by including fitted values of  $\widehat{\ln(vol')}_{ih}$  instead of  $\ln(vol)_{ih}$ . As outlined, we assume  $\beta_{A1}$  in equation (1) to be biased by selective-referral. Hence, after eliminating any effects of selective-referral, a decline in the case volume effect should be observable. However, if unobserved patient heterogeneity is another main driving factor, an increase of the volume coefficient in equation (3) is possible. With the instrumental variable approach it is not possible to separate the effect the selective-referral bias from the unobserved heterogeneity bias (Hockenberry et al., 2010) but, because the two biases work in different directions, it is possible to determine the stronger influence.

---

<sup>9</sup>It is possible to argue that we are not on the hospital level, but on department level for each hospital as we analyze one condition only. Hence, it may be that quality of hospital  $i$  for a specific condition like hip fractures influences the number of further hospitals treating the same condition in the regional area. Some hospitals might be driven out of the market due to outstanding quality of another hospital, i.e. the number of further hospitals treating the same condition might be endogenous as well. However, we think even this is a rather rare event. Nevertheless, as a robustness check we use the number of any further hospital in the regional area of hospital  $i$  as instrument regardless whether they treat hip fracture patients or not. The results do not change.

## 4 Results

We find a strong negative effect of case volume on outcomes ( $p < 0.01$ ) (Table 3). Patients who are treated in hospitals with a higher case volume have a lower probability of death. Our variables adjusting for patient characteristics show the expected sign. Older patients have a higher risk to die, as do male patients. The probability of death increases also for patients with more comorbidities compared to patients who have no comorbidities. The dummies for the CCI show a steady increase of this risk, i.e. the patients with a CCI of zero have the lowest probability of death, followed by patients with one or two comorbidities and so on. The differences between two adjoining Charlson categories are significant ( $p < 0.05$ ). Patients who are admitted during winter time have also a significant higher mortality risk in comparison to patients who are admitted during the other months. We cannot identify a significant weekend effect. Being transferred from another hospital increases the probability of death but being admitted for emergency has no significant effect on mortality compared to patients who have a scheduled admission. Patients with a femoral neck fracture have a higher risk of death compared to patients with a pertrochanteric fracture. Moreover, patients who are transferred during their hospital stay have a higher risk to die compared to patients who stay at the same department. The ownership of a hospital has no influence on mortality. Patients who are treated in a teaching hospital or in hospitals with more beds have a higher probability of death. Patients who are treated in hospitals with an ICU have a lower probability of death compared to patients without an ICU. These results are in line with prior research (Hentschker and Mennicken, 2012).

However, the OLS coefficient does only show a correlation between volume and outcome rather than a causal effect and hence, we turn to the IV estimation results: The first stage regression (Table A3 in the Appendix) shows that the instruments are separately statistically significant ( $p < 0.01$ ) and have also the sign we expect: The more patients with hip fracture are in the regional area of hospital  $i$ , the higher the case volume is and the more hospitals are around the regional area of hospital  $i$ , the lower the case volume of hospital  $i$ . The instruments are also jointly significant with an F-test of 19.7 which is above the general accepted value of 10 (Stock et al., 2002) and hence, problems with weak instruments do not appear in our case.<sup>10</sup> We also apply a test for endogeneity. We have to reject the null hypothesis that case

---

<sup>10</sup>We obtain even an F-value of 42 if we do not control for the county types in our first stage

volume is exogenous at the 5% level and conclude that IV regression is necessary.

The IV coefficient (Table 3) reflects the plain effect of volume on outcome. The coefficient is 1.5 times higher than the OLS coefficient and still highly significant.<sup>11</sup> In quantitative terms, the predicted probability of death decreases by 0.00025 percentage points for every one percentage increase in case volume for every patient. Our patient and hospital characteristics besides case volume stay basically the same in the IV model compared to the OLS model.

Table 3: Regression results

	OLS		IV	
	Coefficient	S.E.	Coefficient	S.E.
In case volume	-0.0169***	(0.0025)	-0.0252***	(0.0050)
Age	0.0025***	(0.0001)	0.0025***	(0.0001)
Male	0.0367***	(0.0022)	0.0367***	(0.0022)
Admission reason: emergency	-0.0013	(0.0021)	-0.0012	(0.0021)
transfer	0.0284***	(0.0077)	0.0280***	(0.0076)
Femoral neck fracture	0.0034**	(0.0016)	0.0032**	(0.0016)
Transfer between departments	0.0401***	(0.0032)	0.0402***	(0.0032)
Winter	0.0053***	(0.0017)	0.0053***	(0.0017)
Weekend	0.0012	(0.0017)	0.0012	(0.0017)
CCI: 1-2	0.0227***	(0.0015)	0.0227***	(0.0015)
3-4	0.0662***	(0.0030)	0.0663***	(0.0030)
≥ 5	0.1369***	(0.0058)	0.1368***	(0.0058)
Ownership: private not-for-profit	-0.0006	(0.0024)	-0.0018	(0.0025)
private for-profit	0.0017	(0.0035)	0.0007	(0.0036)
University hospital	0.0068	(0.0072)	0.0048	(0.0071)
Teaching hospital	0.0054**	(0.0026)	0.0074***	(0.0028)
Beds: 201-499	0.0081***	(0.0031)	0.0114***	(0.0036)
≥ 500	0.0107**	(0.0042)	0.0163***	(0.0051)
ICU	-0.0150***	(0.0023)	-0.0143***	(0.0023)
Constant	-0.1166***	(0.0119)	-0.0842***	(0.0204)
Observations	89,203		89,203	
Number of hospitals	1,155		1,155	

Notes: \* p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01; cluster-robust standard errors computed.

## 5 Robustness checks

When constructing the instruments we used driving time by car as the measure of distance, i.e. we take geographic and infrastructural differences into account. This point is often neglected (e.g. Seider et al., 2004; Tay, 1999), but is especially important for more rural areas because

regression because the F-test is sensitive for the number of variables.

<sup>11</sup>It has been shown that the 2SLS estimator is consistent but biased and that the bias is increasing with the number of instruments (Angrist and Pischke, 2009). An alternative is the limited information maximum likelihood (LIML) estimator which reduces this bias. In our case the results stay the same and we go further with the 2SLS estimation results.

straight-line measurements would underestimate access in regions with less comprehensive infrastructure. Therefore, as a robustness check, we also considered street kilometers and straight-lines as further measures of distance. Hence, we re-estimated equations (1) - (3) using instruments derived from street kilometers and straight-line distances. Furthermore, we varied the arbitrarily set distance thresholds of the instruments discussed in section (2). In Table (4) we show main statistics from equations (1) - (3) to assess how results depend on the choice and the thresholds of the instruments.

The coefficient of case volume shows moderate differences in second stage regression, i.e. the coefficient with straight-line distance is almost 25 % larger than the coefficient with travel time. Furthermore, while the F-test is always above the accepted threshold 10, the respective p-value for testing endogeneity of volume shows substantial variation. If we take a strict threshold of 5% for statistical significance as in Gaynor et al. (2005), we would conclude that volume is exogenous for specifications (1), (4) and (5) when using the Charlson comorbidity index for risk-adjustment. The same specifications using Elixhauser diagnoses tend to have lower p-values. The specification with straight-line distances as instruments always leads to the conclusion of an endogeneity concern.

As summarized in section 1 it is possible to analyze the volume-outcome relationship on the patient as well as on the hospital level. In order to compare results between both levels of analyzes, we estimate the former regressions on the hospital level. To do so, we need to aggregate the data set, i.e. particularly the patient characteristics, on the hospital level. Therefore we estimate equation (1) without hospital characteristics and without case volume using logistic regression (equation (4)). From this equation we calculate the predicted probability of death for each individual and accumulate them per hospital to obtain the expected deaths per hospital. The dependent variable on the hospital level is then the observed expected ratio (O/E ratio), which is above one if the hospital has more deaths than expected (poor quality) or below one if the hospital has fewer deaths than expected (good quality). We formulate our estimation models analogue to the previous approach with the basis OLS regression (equation (5)) and the first stage and second stage equation, equations (6) and (7),

Table 4: Comparison of instruments

Type of instrument	Charlson			Elixhauser		
	Test of endogeneity (p-value)	F-test (instrument relevance)	Case volume IV coefficient (S.E.)	Test of endogeneity (p-value)	F-test (instrument relevance)	Case volume IV coefficient (S.E.)
<b>Potential patients: 0 to 10, 10 to 20, 20 to 30; number of further hospitals: 0 to 15, 15 to 30</b>						
(1) travel time in minutes	0.051	19.715	-0.0252*** (0.0050)	0.007	19.865	-0.0296*** (0.0051)
(2) distance in km	0.023	17.233	-0.0269*** (0.0053)	0.005	17.313	-0.0305*** (0.0054)
(3) straight-line distance in km	0.003	14.873	-0.0312*** (0.0056)	0.000	15.001	-0.0359*** (0.0057)
<b>Potential patients: 0 to 5, 5 to 10, 10 to 15, 15 to 20; number of further hospitals: 0 to 15, 15 to 30</b>						
(4) travel time in minutes	0.084	19.064	-0.0245*** (0.0051)	0.012	19.147	-0.0464*** (0.0056)
(5) distance in km	0.067	15.887	-0.0252*** (0.0052)	0.011	15.955	-0.0293*** (0.0053)
(6) straight-line distance in km	0.002	15.158	-0.0315*** (0.0055)	0.000	15.287	-0.0296*** (0.0053)

Notes: \* p &lt; 0.10, \*\* p &lt; 0.05, \*\*\* p &lt; 0.01.



respectively.

$$y_i = \alpha_0 + \mathbf{x}'_i \boldsymbol{\beta}_{C1} + \mu_i \quad (4)$$

$$O/E_h = \alpha_0 + \ln(vol')_h \beta_{D1} + \mathbf{k}'_h \boldsymbol{\beta}_{D2} + \eta_h \quad (5)$$

$$\ln(vol)_h = \alpha_0 + \mathbf{k}'_h \boldsymbol{\pi}_{B2} + \mathbf{p}'_h \boldsymbol{\gamma}_{B1} + \mathbf{h}'_h \boldsymbol{\gamma}_{B2} + \mathbf{s}'_h \boldsymbol{\gamma}_{B3} + \lambda_h \quad (6)$$

$$O/E_h = \alpha_0 + \widehat{\ln(vol')}_h \beta_{E1} + \mathbf{k}'_h \boldsymbol{\beta}_{E2} + \kappa_h \quad (7)$$

Table 5 shows an extract from our main results.<sup>12</sup> Basically the results stay the same. The case volume coefficients show the same sign and direction as in the patient level analysis. The size of the coefficient differs, but this is because of different dependent variables and hence different interpretation levels. We also estimate the above regressions and exclude every hospital which has less than one expected death. This is recommended for statistical reasons (Ash et al., 2003) and again, results do not change.

Table 5: Estimation results on hospital level

	<b>Charlson</b>	<b>Elixhauser</b>
In case volume		
OLS coefficient (S.E.)	-0.1586*** (-0.0398)	-0.1252*** (-0.039)
IV coefficient (S.E.)	-0.4695*** (-0.1052)	-0.4239*** (-0.103)
Estimation with hospital characteristics	Yes	Yes
F-test (instrument relevance)	15.304	15.304
Test of endogeneity (p-value)	0.001	0.001
Observations	1,155	1,155

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 6 Discussion

The volume-outcome relationship can be explained by two hypotheses which have reverse causality directions: the practice-makes-perfect hypothesis and the selective-referral hypothesis. The first hypothesis states that through learning effects hospitals obtain a good quality. The second hypothesis asserts that quality is the reason for a high case volume. To determine only the effect of practice-makes-perfect it is necessary to exclude selective-referral with an instrumental variable approach. After exclusion a decline in the volume coefficient is expected because the effect of having a high case volume due to a good quality is excluded. But in

<sup>12</sup>Table A5 in the Appendix shows descriptive statistics on the hospital level.

modeling the volume-outcome relationship also the aspect of unobserved patient heterogeneity has to be considered. If we cannot sufficiently control for patient characteristics, the error term is likely to be correlated with quality and case volume, i.e. unobserved sicker patients choose or are referred to hospitals providing better quality. These unobserved sicker patients have also a higher risk of death. Hence, excluding unobserved heterogeneity would yield to an increase in the case volume coefficient. Following these arguments the pure exclusion of selective-referral yields to a decline in the case volume coefficient, but the combination of selective-referral with unobserved patient heterogeneity could yield to an increase in the case volume coefficient if the latter effect exceeds the selective-referral effect. This is what we observe in the presented estimations. The interpretation stays the same when we estimate the models on the hospital level. Hence, at least in this analysis it does not matter if patient or hospital level data is used.

There are some limitations regarding the data. We have only a cross section available and, therefore, can only show differences between hospitals. With panel data and enough variation in case volume, it would have been possible to show effects within hospitals and rule out unobserved heterogeneity which is constant over time with fixed effects. However, the included hospital characteristics are predominantly significant and should therefore capture these effects. Furthermore, we have no information about staffing levels and accordingly, the case numbers of involved surgeons are missing. Hence, we can only use the case volume per hospital. We argue that for obtaining a good quality in a hospital the team process is the key factor.

## 7 Conclusion

In this paper we find further evidence for a causal relationship from volume to outcome. We are the first to provide evidence for the practice-makes-perfect hypothesis using German data. After eliminating the effects of selective-referral in the IV regression, we observe an increase of about 50% in the coefficient of volume. These results are in line with Seider et al. (2004) and accordingly Gaynor et al. (2005), who also present results with an increase in the volume coefficient after using IV. However, both did not give any attention to this result and only stated that the coefficient of volume should decrease when eliminating selective-

referral. Gaynor et al. (2005) rejected the null hypothesis of no differences between OLS and IV regression with a p-value of 0.06 and concluded that volume is exogenous. However, results in Table 4 show that the p-value is sensitive to the specific type of instrument used. Hence, we recommend a more conservative approach of using a threshold of 0.10 for testing.

Based on the theoretical outline, the coefficient of volume should decrease after eliminating selective-referral effects. So far, the international literature is lacking explanations for the observed, but counterintuitive increase of the volume effect when controlling for selective-referral. We argue that this downward bias of the OLS coefficient is caused by unobserved patient heterogeneity.

We could show that the effect of volume is even stronger than in normal OLS regressions. Hence, there is a possibility that minimum volume standards can achieve the goal of quality improvement in the hospital sector, at least for the treatment of hip fracture patients. But before its introduction more analyzes are needed for deriving adequate thresholds and assessing changes in access to hospital services in different regions.

# A Appendix

Table A1: Definition of settlement structures (SST)

Abbr.	Contextual type	Differentiated county type	Characteristics
SST 1	Agglomerated regions	Core cities	Cities without county membership > 100,000 inhabitants
SST 2		Very densely populated counties	Counties with more than 300 inhabitants/km <sup>2</sup>
SST 3		Densely populated counties	Counties with more than 150 inhabitants/km <sup>2</sup>
SST 4		Rural counties	Counties with less than 150 inhabitants/km <sup>2</sup>
SST 5	Urbanized regions	Core cities	Cities without county membership > 100,000 inhabitants
SST 6		Densely populated counties	Counties with more than 150 inhabitants/km <sup>2</sup>
SST 7		Rural counties	Counties with less than 150 inhabitants/km <sup>2</sup>
SST 8	Rural regions	Densely populated rural counties	Counties with more than 100 inhabitants/km <sup>2</sup>
SST 9		Sparsely populated rural counties	Counties with less than 100 inhabitants/km <sup>2</sup>

Source: BBSR - Bundesinstitut für Bau Stadt- und Raumforschung (2012).

Table A2: Probit Model and IV-Probit

	<b>Probit</b>		<b>IV-Probit</b>	
	Coefficient	S.E.	Coefficient	S.E.
ln case volume	-0.1494***	(0.0217)	-0.2315***	(0.0465)
Age	0.0320***	(0.0009)	0.0321***	(0.0009)
Male	0.3071***	(0.0167)	0.3071***	(0.0166)
Admission reason: emergency	-0.0074	(0.0187)	-0.0064	(0.0187)
transfer	0.2246***	(0.0548)	0.2206***	(0.0539)
Femoral neck fracture	0.0202	(0.0140)	0.0181	(0.0141)
Transfer between departments	0.3331***	(0.0227)	0.3342***	(0.0224)
Winter	0.0435***	(0.0148)	0.0425***	(0.0148)
Weekend	0.0065	(0.0153)	0.0069	(0.0153)
CCI: 1-2	0.3532***	(0.0198)	0.3530***	(0.0198)
3-4	0.6375***	(0.0233)	0.6375***	(0.0232)
≥ 5	0.9705***	(0.0288)	0.9689***	(0.0288)
Ownership: private not-for-profit	-0.0119	(0.0209)	-0.0246	(0.0219)
private for-profit	0.0130	(0.0304)	0.0029	(0.0315)
University hospital	0.0489	(0.0651)	0.0275	(0.0643)
Teaching hospital	0.0387*	(0.0227)	0.0588**	(0.0244)
Beds: 201-499	0.0710***	(0.0273)	0.1035***	(0.0318)
≥ 500	0.1003***	(0.0370)	0.1548***	(0.0453)
ICU	-0.1553***	(0.0215)	-0.1471***	(0.0218)
Constant	-4.1205***	(0.1200)	-3.7961***	(0.2023)
Observations	89,203		89,203	
Number of hospitals	1,155		1,155	

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ; cluster-robust standard errors computed.

Table A3: First stage regression explaining case volume

	IV	
	Coefficient	S.E.
pot_pat_0to10	0.0027***	(0.0003)
pot_pat_10to20	0.0006***	(0.0001)
pot_pat_20to30	0.0003***	(0.0001)
hos_0to15	-0.1154***	(0.0094)
hos_15to30	-0.0216***	(0.0044)
County type 2	0.0114	(0.0518)
County type 3	0.0344	(0.0625)
County type 4	-0.1689**	(0.0766)
County type 5	0.1579**	(0.0783)
County type 6	0.0714	(0.0631)
County type 7	-0.0538	(0.0674)
County type 8	0.0902	(0.0664)
County type 9	-0.0534	(0.0791)
Age	0.0006***	(0.0002)
Male	0.0007	(0.0033)
Admission reason: emergency	0.0199	(0.0138)
transfer	-0.0234	(0.0332)
Femoral neck fracture	-0.0224***	(0.0036)
Transfer between departments	0.0256	(0.0163)
Winter	-0.0067**	(0.0032)
Weekend	0.0043	(0.0031)
CCI: 1-2	0.0041	(0.0048)
3-4	0.0037	(0.0075)
≥ 5	-0.0053	(0.0100)
Ownership: private not-for-profit	-0.1134***	(0.0282)
private for-profit	-0.1515***	(0.0410)
University hospital	-0.2108**	(0.0901)
Teaching hospital	0.1649***	(0.0306)
Beds: 201-499	0.3722***	(0.0366)
≥ 500	0.6004***	(0.0496)
ICU	0.0066	(0.0311)
Constant	3.8605***	(0.0732)
Observations	89,203	
Number of hospitals	1,155	

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ;  
cluster-robust standard errors computed.

Table A4: Reduced form regression

	<b>IV</b>	
	Coefficient	S.E.
pot_pat_0to10	-0.0001***	(0.0000)
pot_pat_10to20	-0.0000**	(0.0000)
pot_pat_20to30	-0.0000**	(0.0000)
hos_0to15	0.0021***	(0.0008)
hos_15to30	0.0012***	(0.0004)
County type 2	-0.0045	(0.0046)
County type 3	0.0001	(0.0061)
County type 4	0.0076	(0.0086)
County type 5	-0.0034	(0.0058)
County type 6	-0.0095*	(0.0056)
County type 7	-0.0060	(0.0061)
County type 8	-0.0032	(0.0059)
County type 9	-0.0042	(0.0067)
Age	0.0025***	(0.0001)
Male	0.0367***	(0.0022)
Admission reason: emergency	-0.0021	(0.0022)
transfer	0.0279***	(0.0078)
Femoral neck fracture	0.0037**	(0.0016)
Transfer between departments	0.0401***	(0.0033)
Winter	0.0054***	(0.0017)
Weekend	0.0011	(0.0017)
CCI: 1-2	0.0225***	(0.0015)
3-4	0.0661***	(0.0030)
≥5	0.1369***	(0.0058)
Ownership: private not-for-profit	0.0003	(0.0026)
private for-profit	0.0045	(0.0035)
University hospital	0.0101	(0.0075)
Teaching hospital	0.0035	(0.0026)
Beds: 201-499	0.0008	(0.0029)
≥ 500	-0.0007	(0.0039)
ICU	-0.0142***	(0.0025)
Constant	-0.1768***	(0.0084)
Observations	89,203	
Number of hospitals	1,155	

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ;  
cluster-robust standard errors computed.

Table A5: Descriptive Statistics on hospital level

	<b>Mean</b>	<b>S.D.</b>
Observed-expected ratio	1.110	0.727
ln case volume	4.157	0.644
Ownership		
Public	0.437	0.496
Private non-profit	0.421	0.494
Private for-profit	0.142	0.349
University hospital	0.033	0.178
Teaching hospital	0.423	0.494
Beds		
200 beds	0.322	0.467
201-499 beds	0.466	0.499
$\geq$ 500 beds	0.212	0.409
ICU	0.366	0.482
Number of hospitals	1,155	



## References

- Agency for Healthcare Research and Quality (2007). *AHRQ quality indicators - Guide to inpatient quality indicators: Quality of care in hospitals - volume, mortality, and utilization, Version 3.1*, Rockville.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, Princeton.
- Ash, A. S., Shwartz, M. and Peköz, E. A. (2003). Comparing outcomes across providers, in L. I. Iezzoni (ed.), *Risk adjustment for measuring health care outcomes*, Health Administration Press, Chicago, pp. 297–333.
- Ayanian, J. Z. and Weissman, J. S. (2002). Teaching hospitals and quality of care: a review of the literature, *The Milbank Quarterly* **80**(3): 569–593.
- Barker, D., Rosenthal, G. and Cram, P. (2011). Simultaneous relationships between procedure volume and mortality: do they bias studies of mortality at specialty hospitals?, *Health Economics* **20**(5): 505–518.
- BBSR - Bundesinstitut für Bau Stadt- und Raumforschung (2012). Indikatoren und Karten zur Raum- und Stadtentwicklung. INKAR. Ausgabe 2012. Referenzdatei Kreise-Raumordnungsregionen-Regionstypen.
- Becker, D. J. (2007). Do hospitals provide lower quality care on weekends?, *Health Services Research* **42**(4): 1589–1612.
- Bell, C. M. and Redelmeier, D. A. (2001). Mortality among patients admitted to hospitals on weekends as compared with weekdays, *New England Journal of Medicine* **345**(9): 663–668.
- Bundesgeschäftsstelle Qualitätssicherung (2008). *BQS-Bundesauswertung 2007, Hüftgelenknahe Femurfraktur*, Düsseldorf.
- Charlson, M. E., Pompei, P., Ales, K. L. and MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation, *Journal of Chronic Diseases* **40**(5): 373–383.
- Elixhauser, A., Steiner, C., Harris, D. R. and Coffey, R. M. (1998). Comorbidity measures for use with administrative data, *Medical Care* **36**(1): 8–27.
- Farley, D. E. and Ozminkowski, R. J. (1992). Volume-outcome relationships and in-hospital mortality: the effect of changes in volume over time, *Medical Care* **30**(1): 77–94.
- Gandjour, A., Bannenberg, A. and Lauterbach, K. W. (2003). Threshold volumes associated with higher survival in health care: a systematic review, *Medical Care* **41**(10): 1129–1141.
- Gaynor, M., Seider, H. and Vogt, W. B. (2005). The volume-outcome effect, scale economies, and learning-by-doing, *The American Economic Review* **95**(2): 243–247.
- Gowrisankaran, G. and Town, R. J. (1999). Estimating the quality of care in hospitals using instrumental variables, *Journal of Health Economics* **18**(6): 747–767.
- Halm, E. A., Lee, C. and Chassin, M. R. (2002). Is volume related to outcome in health care? A systematic review and methodologic critique of the literature, *Annals of Internal Medicine* **137**(6): 511–520.
- Hamilton, B. H. and Hamilton, V. H. (1997). Estimating surgical volume-outcome relationships applying survival models: accounting for frailty and hospital fixed effects, *Health Economics* **6**(4): 383–395.
- Hentschker, C. and Mennicken, R. (2012). The relationship between quality and hospital case volume: an empirical examination with German data, *Ruhr Economic Paper* (No. 341).

- Hockenberry, J. M., Lien, H.-M. and Chou, S.-Y. (2010). Surgeon and hospital volume as quality indicators for CABG in Taiwan: examining hazard to mortality and accounting for unobserved heterogeneity, *Health Services Research* **45**(5 Pt 1): 1168–1187.
- Iezzoni, L. I. (2003). Reasons for risks adjustment, in L. I. Iezzoni (ed.), *Risk adjustment for measuring health care outcomes*, Health Administration Press, Chicago, pp. 1–16.
- Kessler, D. P. and McClellan, M. B. (2000). Is hospital competition socially wasteful?, *The Quarterly Journal of Economics* **115**(2): 577–615.
- Luft, H. S., Bunker, J. P. and Enthoven, A. C. (1979). Should operations be regionalized? The empirical relation between surgical volume and mortality, *The New England Journal of Medicine* **301**(25): 1364–1369.
- Luft, H. S., Hunt, S. S. and Maerki, S. C. (1987). The volume-outcome relationship: practice-makes-perfect or selective-referral patterns?, *Health Services Research* **22**(2): 157–182.
- Milcent, C. (2005). Hospital ownership, reimbursement systems and mortality rates, *Health Economics* **14**(11): 1151–1168.
- Norton, E. C., Garfinkel, S. A., McQuay, L. J., Heck, D. A., Wright, J. G., Dittus, R. and Lubitz, R. M. (1998). The effect of hospital volume on the in-hospital complication rate in knee replacement patients, *Health Services Research* **33**(5 Pt 1): 1191–1210.
- Ozimek, A. and Miles, D. (2011). Stata utilities for geocoding and generating travel time and travel distance information, *Stata Journal* **11**(1): 106–119.
- Quan, H., Sundararajan, V., Halfon, P., Fong, A., Burnand, B., Luthi, J.-C., Saunders, L. D., Beck, C. A., Feasby, T. E. and Ghali, W. A. (2005). Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data, *Medical Care* **43**(11): 1130–1139.
- Seider, H., Gaynor, M. and Vogt, W. B. (2004). Volume-outcome and antitrust in US health care markets, *unpublished working paper*.
- Stock, J. H., Wright, J. H. and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments, *Journal of Business & Economic Statistics* **20**(4): 518–529.
- Tay, A. (1999). Measuring hospital quality, *Working Paper*.
- Tsai, A. C., Votruba, M., Bridges, J. F. P. and Cebul, R. D. (2006). Overcoming bias in estimating the volume-outcome relationship, *Health Services Research* **41**(1): 252–264.