RUHR
ECONOMIC PAPERS

Corinna Hentschker
Roman Mennicken

Selective-referral and Unobserved Patient Heterogeneity – Bias in the Volume-outcome Relationship

#527

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #527

Responsible Editor: Thomas Bauer

ISSN 1864-4872 (online) – ISBN 978-3-86788-602-4

Ruhr Economic Papers #527

Corinna Hentschker and Roman Mennicken

Selective-referral and Unobserved Patient Heterogeneity – Bias in the Volume-outcome Relationship

RUHR UNIVERSITÄT BOCHUM RUB

rwi

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788602
ISSN 1864-4872 (online)
ISBN 978-3-86788-602-4

Corinna Hentschker and Roman Mennicken[1]

Selective-referral and Unobserved Patient Heterogeneity – Bias in the Volume-outcome Relationship

Abstract

This paper examines the causal effect of the experience of a hospital with treating hip fractures (volume) on treatment outcome for patients. A full sample of administrative data from Germany for the year 2007 is used. We apply an instrumental variable approach to eliminate endogeneity concerns due to reverse causality and unobserved patient heterogeneity. As instruments for case volume we use the number of potential patients and the number of further hospitals in the region around every hospital. Our results indicate that after application of an IV regression of volume on outcome, volume significantly increases quality.

JEL Classification: I11, I12, I18

Keywords: volume; hospital quality; mortality; instrumental variables

November 2014

1 Corinna Hentschker, RWI and RUB; Roman Mennicken, Landschaftsverband Rheinland, Köln. *– We thank Klaus Focke, Uwe Mehlhorn and Daniel Viehweg from the BKK Federal Association. Furthermore, we thank Thomas Bauer, Martin Fischer, Harald Tauchmann, Colin Vance, and Ansgar Wübker for helpful remarks as well as Rüdiger Budde for his support with geocoding of our data. Moreover, we are grateful for comments at the ECHE, CINCH Academy, and Verein für Socialpolitik 2014. The administrative data of §21 KHEntgG was used as part of a cooperation agreement for the further development of the DRG-system from April, 1st 2011 between the RWI and the BKK Federal Association. The opinions expressed in this article are the authors' own and do not necessarily reflect the views of the affiliated institutions.* *– All correspondence to Corinna Hentschker, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: corinna.hentschker@rwi-essen.de*

1 Introduction

The volume-outcome relationship defines the relationship between case volume and quality of a hospital. In 1979, Luft et al. (1979) showed in their seminal paper that in ten out of twelve procedures a correlation between volume and outcome exists. Subsequent studies of this relationship predominantly confirmed this correlation (Gandjour et al., 2003; Halm et al., 2002). However, these studies inherently assume that volume is the driving factor for outcome (practice-makes-perfect hypothesis). Nevertheless, the relationship between volume and outcome does not necessarily reflect a causal effect. Volume might be endogenous because of reverse causality, i.e. outcome determines volume (selective-referral hypothesis), or because unobserved characteristics drive this relationship, e.g. unobserved sicker patients choose hospitals with a higher case volume.

The necessity for assessing the causal effect of volume on outcome arises from its political impact: The volume-outcome effect is the foundation for minimum volume standards. For example, in Germany minimum volume standards were introduced for five interventions in 2004. Consequently, hospitals which do not achieve a certain number of cases within a specific diagnosis are not allowed to treat patients with this diagnosis anymore. This regulation was introduced following international evidence confirming a positive relationship between volume and outcome. However, from a health policy point of view, the causal direction in volume-outcome matters as minimum volume standards 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 trade-off between potential gains in quality of care and losses in access to care. If the practice-makes-perfect hypothesis indeed 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, 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 referral effect, the causality has to be determined.

The recent literature is quite sparse in determining the causal effect of the volume-outcome relationship. Only a few studies use instrumental variable regression or simultaneous equation models to overcome endogeneity. Common instruments are the number of hospital beds (Allareddy et al., 2012; Farley and Ozminkowski, 1992; Luft et al., 1987; Norton et al., 1998) and geographical factors (Barker et al., 2011; Seider et al., 2004; Tsai et al., 2006). The number of hospital beds and hence the size of a hospital, however, has been shown to influence quality directly (Keeler et al., 1992), i.e. larger hospitals show a better quality. Therefore, this instrument may be invalid. Avdic et al. (2014) use closure of hospitals as instrument. The closure of hospitals might not be a random exogenous shock. Hospitals may close rather due to quality concerns, rendering this instrument also to be invalid. Hamilton and Hamilton (1997) use a duration model with hospital fixed effects. However, reverse causality can still drive their results because the authors only exclude fixed quality differences over time with hospital fixed effects, but no time-varying differences. Furthermore, fixed effects cannot solve the problem of reverse causality.

This paper examines the causal relationship of volume on outcome for patients with hip fracture. Hip fractures are a common reason for hospital admission of the elderly, with a comparatively high mortality rate. Therefore, it is important to detect factors that drive quality of treatment. We extend the analysis of Hentschker and Mennicken (2014), who showed a correlation of volume and outcome for hip fracture patients. We use an instrument similar to Seider et al. (2004) and Gaynor et al. (2005)[1]. The instruments are the number of potential patients and the number of further hospitals in the region around every hospital. The case volume of a hospital should increase if more patients with hip fracture live around the hospital and decrease if further hospitals treat patients with hip fracture in the regional area around the hospital. Using administrative data from all inpatients in Germany in 2007, we find that after application of instrumental variable regression volume has a significant positive effect on outcome.

Our analysis contributes to the literature in three ways. First, to our knowledge we are the first who provide empirical evidence for a causal volume-outcome relationship for Germany. This is of particular importance because minimum volume standards have already been introduced in Germany without considering the causal relationship. Second, we use a full sample of all inpatients available and not only for a specific group, e.g. Medicare patients as Barker et al. (2011) and Tsai et al. (2006). Therefore, we are able to determine the real case

[1] Gaynor et al. (2005) is the corresponding published article of the working paper of Seider et al. (2004).

volume of a hospital and not only the case volume of a subgroup of patients where the share of patients can differ between hospitals. Third, the majority of the causal volume-outcome literature focuses only on the bias through reversed causality (Barker et al., 2011; Farley and Ozminkowski, 1992; Gaynor et al., 2005). We describe the biases through selective-referral and unobserved patient heterogeneity in detail and explain the different directions of the distortions.

The remainder of this paper is organized as follows: Section 2 explains the endogeneity concerns when investigating the volume-outcome relationship and presents the empirical strategy. Section 3 describes the data and variables used in the analysis. Section 4 shows the estimation results. Section 5 summarizes the main findings and concludes.

2 Empirical Strategy

We specify our dependent variable y_{ihc} as a binary variable, which indicates whether patient i has died in hospital h in county type c, where the latter are distinguished by population density, after being treated because of a hip fracture. We estimate the following regression model via OLS[2]:

$$y_{ihc} = \alpha_0 + \beta_1 ln(vol)_{hc} + \boldsymbol{x'_{ihc}\beta_2} + \boldsymbol{k'_{hc}\beta_3} + \boldsymbol{s'_c\beta_4} + \varepsilon_{ihc} \tag{1}$$

where $ln(vol)_{hc}$ is the logarithm of case volume, x_{ihc} are patient characteristics, k_{hc} are hospital characteristics, s_c are county type indicators and ε_{ihc} is a random error term. The coefficient β_1 is of primary interest; it measures how case volume effects the outcome of a hospital.

Regression model (1) neglects possible endogeneity threats. Volume can be endogenous for two reasons: reverse causality and unobserved patient heterogeneity. Reverse causality may occur if higher quality results in a higher volume rather than a higher volume to better quality. The practice-makes-perfect hypothesis states that higher case volume leads to a better quality because of learning effects, and economies of scale (Luft et al., 1987; Seider et al., 2004).

[2]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 A4). The marginal effects of the probit regression are almost identical to the average marginal effects of the LPM and IV regression.

Hospitals that treat more patients with a specific condition reduce their mistakes, optimize processes, and develop better routines. 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 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 with the lowest mortality rate. Based on these arguments quality is the leading cause of a high case volume. Both hypotheses are possible resulting in OLS estimates of β_1 to be biased downwards. Further, the volume-outcome relationship might be biased by an omitted variable bias due to unobserved patient heterogeneity. Patients may choose hospitals based on their initial health status (Tay, 1999). Patient characteristics are usually unequally distributed across hospitals. University hospitals, for example, often treat sicker patients in terms of age and comorbidities. If the health status of the patients is not fully observed, unobserved patient characteristics captured in the error term may be correlated both with the volume variable and the outcome variable (Iezzoni, 2003). Therefore, information on patient characteristics is essential to control for in order to adequately identify the volume-outcome relationship.

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. 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. Unobserved characteristics would lead to an upward bias of the effect of volume in a regression of volume on outcome.

We use an instrumental variable approach to correct for these endogeneity problems. To implement this strategy, we require an instrument Z that is strongly correlated with volume ($cov(ln(vol)_{hc}, Z_{hc}) \neq 0$) but uncorrelated with the error term ($cov(Z_{hc}, \varepsilon_{ihc}) = 0$). We use two instruments similar to Seider et al. (2004). In general, patients choose hospitals that are closer to their residence. This implies that the case volume of a hospital depends on the number of potential patients p_{hc} and the number of further hospitals h_{hc} in the region around every hospital which we are using as instruments (see section 3). 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 respective region.

The number of potential patients and the distance of each patient to a hospital should have no direct influence on the quality of treatment. Patients' residences can be considered as exogenous to hospital quality, because it is unlikely that patients choose residency on the basis of quality of care in a nearby hospital. There might be other unobserved factors that are correlated with patient's residence and outcome of a hospital, e.g. income. These differences should be captured by county type indicators. It is also possible that the number of further hospitals in the regional area may be influenced by the quality of treatment of a nearby hospital. For example some hospitals might be driven out of the market due to the outstanding quality of another hospital, i.e. the number of further hospitals treating the same condition might be endogenous as well. To avoid these possible endogeneity threats we control for concentration in the hip fracture market using the Herfindahl Hirschman Index (HHI). Conditioned on the HHI, the number of further hospitals should not have a direct influence on quality. Population and hospital density show substantial variation throughout Germany. To take these differences into account we use county type indicators.

With these instruments we specify the following first-stage equation (2), where the logarithm of case volume is regressed on all covariates of equation (1) and the instruments p_{hc} and h_{hc}. In the second-stage equation (3) the fitted values of $ln(vol)_{ihc}$ from equation (2) are used to model the causal effect of volume on outcome.

$$ln(vol)_{ihc} = \alpha_0 + \boldsymbol{x'_{ihc}\pi_1} + \boldsymbol{k'_{hc}\pi_2} + \boldsymbol{s'_{c}\pi_3} + \boldsymbol{p'_{hc}}\gamma_1 + \boldsymbol{h'_{hc}}\gamma_2 + \upsilon_{ihc} \tag{2}$$

$$y_{ihc} = \alpha_0 + \beta_1 \widehat{ln(vol)}_{ihc} + \boldsymbol{x'_{ihc}\beta_2} + \boldsymbol{k'_{hc}\beta_3} + \boldsymbol{s'_{c}\beta_4} + \nu_{ihc} \tag{3}$$

For the estimation we use the two-step generalized method of moment (GMM) IV estimator (IV-GMM).[3] Note that the IV approach in equation (3) identifies only a local average treatment effect (LATE), i.e. we only measure the effect for hospitals that are influenced by the number of potential patients and further hospitals in the regional area (compliers).

[3]The IV-GMM estimator is a more general framework than the two-stage least square (2sls) IV estimator. To simplify the representation of the formulas we choose the form of the 2sls approach, keeping in mind that we estimate with GMM.

3 Data

We use administrative data of all German hospitals for the year 2007. The data contains the total in-patient population except psychiatric cases in Germany. It provides detailed information for the patients 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 location. We geo-coded the addresses of hospitals and the centroids of all German ZIP codes[4], so we are able to calculate the distance for each patient to the hospital and distances between hospitals (Ozimek and Miles, 2011).

In the empirical analysis 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 Qualiätssicherung, 2008). We only include patients with a main diagnosis of HIP and a matching procedure code. We exclude 21 patients with missing patient characteristics and 821 patients because they have no valid zip-code. For the last group distances are not computable which we need for the construction of the instruments. We leave out 133 patients who are younger than 20 years because those patients might need a special treatment compared to older patients. Furthermore, we drop patients who have a recorded discharge reason transfer to another hospital (n = 9,210). For those patients we are not able to determine the outcome of the treatment. Our final sample consists of 89,541 patients treated in 1,238 hospitals.[5]

Generally, our data allows us to identify each hospital by a unique identifier. Using data from other sources, we have been able to identify for 47 hospitals further hospital locations under the same identifier, i.e. the patients were treated in 81 hospital locations but the original data set can only distinguish 47 hospitals. For those hospitals, we are not able to identify the actual location, the patient was treated. In this case, we have randomly assigned patients into possible hospital locations based on the share of hip fracture patients that has been documented in the quality reports.[6] This yields to the already mentioned 1,238 hospitals.

[4]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.

[5]The final sample differs from the sample in Hentschker and Mennicken (2014). In the mentioned paper hospitals with less than 10 cases were excluded. We included hospitals with less than 10 cases. The estimation results, which are available from the authors upon request, do not change.

[6]We estimate all models without division of the sample. The results stay basically the same and

We use in-hospital mortality[7] as outcome measure. Mortality is the most frequently used outcome measure in volume-outcome studies for two reasons: First, mortality is a clear defined outcome. This is of importance because every hospital records its own data and coding differences might apply between hospitals but are impossible for mortality. Second, 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. In our model we use the logarithm of volume (Farley and Ozminkowski, 1992; Hamilton and Hamilton, 1997).

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 2) and are similar to Seider et al. (2004). For the number of potential patients we specify three variables, i.e. we choose three radii 0 to 10 minutes, 10 to 20 minutes and 20 to 30 minutes, and sum up all patients with hip fracture that reside within these radii, irrespective of whether they have been treated in the hospital or not. Because of the lower number of hospitals we specify this instrument with only two variables and choose two radii, i.e. 0 to 15 minutes and between 15 and 30 minutes, and sum up every hospital that treats hip fractures. The thresholds of the radii are somewhat arbitrary. We estimated models with different thresholds but the results basically do not change.

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 Elixhauser comorbidities 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 Elixhauser comorbidities are frequently used to do risk adjustment with administrative data (Elixhauser et al., 1998).[8] They consist of 30 diagnoses which are not directly related to the main diagnosis but potentially increase the probability of a worse

are available upon request.

[7] Unfortunately, we are not able to track patients after discharge with the data available. Hence, we cannot consider out-of-hospital mortality.

[8] Another well established specification for risk-adjustment is the Charlson comorbidity index (CCI) (Charlson et al., 1987). We estimated our models using both methods. The results are similar. Regression results with CCI can be found in the Appendix in Table A3.

outcome compared to a patient without such a diagnosis, e.g. diabetes, congestive heart failure and hypertension (Elixhauser et al., 1998). We specify each diagnosis as binary variable that is 1 if the patient has this illness and 0 otherwise. For this purpose, we use diagnosis codes developed by Quan et al. (2005) who mapped the original codes from ICD-9 system to the ICD-10 system used in Germany. Furthermore, we add dummy variables for admission during winter time[9] as well as for admissions over weekend and public holidays. The first variable captures possible seasonal patterns during winter time because people slip on icy grounds. The latter variable captures weekend and public holiday effects due to lower staffing levels in comparison to weekdays (Bell and Redelmeier, 2001; Kuntz et al., 2014). We further include a binary variable that differentiates patients with a femoral neck fracture and a pertrochanteric fracture. As another covariate we use a binary variable that 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), teaching status (Ayanian and Weissman, 2002), and size of a hospital (Keeler et al., 1992) 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. To take the possible influence of competition between hospitals and the quality of treatment into account, we calculate the HHI as in Hentschker et al. (2014).[10] The HHI is a measure of market concentration and can range from 0 to 1.

We further use county type indicators to differentiate hospitals that are located in areas with different population densities as well as urban and rural areas. The indicators range from cities without county membership with more than 100,000 inhabitants to counties with less than 100 inhabitants per square kilometer. An overview of the indicators can be found in Schmid and Ulrich (2013).

Table 1 shows descriptive statistics of the patient and hospital characteristics.[11] On average 6.3% of the patients die in hospital. Most of the patients were female (75%) and

[9]This variable is 1, if the admission was in the months November, December, January, or February and 0 otherwise.

[10]For the calculation of the HHI, product market and geographic market have to be defined. We define the product market based on the single diagnosis; in our case we only include hospitals in the calculation of the HHI which treat hip fracture patients. For the definition of the geographic market we use the 60/01-rule as in in Hentschker et al. (2014).

[11]The descriptive statistic for the distribution of the Elixhauser comorbidities are shown in the Appendix in Table A1.

on average 80 years old. 28% of the patients were admitted on a weekend and 33% during winter time. 42% of the hospitals which treat hip fracture were public hospitals, 41% teaching hospitals, and 3% university hospitals. On average 82 patients with hip fracture live around a hospital within a radius of 10 minutes. Two additional hospitals are within a radius of 15 minutes around each hospitals.

Table 1: Descriptive statistics of HIP

	Mean	**S.D.**	**Min**	**Max**
Patient level (N = 89,541)				
Death	0.063	0.243	0	1
Age	79.546	11.299	20	108
Male	0.251	0.434	0	1
Emergency	0.761	0.427	0	1
Transfer	0.020	0.141	0	1
Femoral neck fracture	0.537	0.499	0	1
Transfer between departments	0.262	0.439	0	1
Winter	0.331	0.471	0	1
Weekend	0.284	0.451	0	1
Hospital level (N = 1,238)				
Case volume	72.327	49.622	1	387
Ownership: private not-for-profit	0.417	0.493	0	1
private for-profit	0.158	0.365	0	1
University hospital	0.032	0.175	0	1
Teaching hospital	0.406	0.491	0	1
Beds: 201-499	0.449	0.498	0	1
$\geq$ 500	0.200	0.400	0	1
ICU	0.347	0.476	0	1
HHI	0.427	0.224	0.057	0.929
Potential patients between 0 to 10 min	81.743	90.771	0	490
10 to 20 min	291.339	366.389	0	1978
20 to 30 min	560.476	647.223	0	3372
Further hospitals between 0 to 15 min	2.354	3.717	0	21
15 to 30 min	10.330	12.909	0	68

4 Results

Table 2 shows the coefficients of ln case volume. Models (1) to (4) show the results of a linear probability model. In the first model we just estimate a bivariate specification of outcome on case volume. We successively add patient characteristics and hospital characteristics in model (2) and (3), respectively. In model (4) we additionally control for county type indicators. We find a strong negative effect of case volume on outcome ($p < 0.01$). Patients who are treated in hospitals with a higher case volume have a lower probability of death. The coefficient increases if patient characteristics and hospital characteristics are added to the model. This indicates that these characteristics are correlated with case volume and are unequally distributed between hospitals with different case volume. Our further control variables show the expected sign. The complete regression results are shown in the Appendix in Table A2.

The LPM coefficients only show a correlation between volume and outcome rather than a causal effect. Therefore, we turn to the IV estimation results. The first-stage regression (Table A5 in the Appendix) shows that the instruments are separately statistically significant ($p < 0.01$) and have the expected sign: The more patients with hip fracture are in the regional area of hospital i, the higher the case volume and the more hospitals are around the regional area of hospital i, the lower the case volume of hospital i. The coefficients decrease with increasing radii. The instruments are also jointly significant with an F-test of 46.8 for the full model which is above the general accepted value of 10 or rather 15 in the case of five instruments (Stock et al., 2002). Hence, problems with weak instruments do not appear in our case. Using the GMM C-statistic, we have to reject the null hypothesis that case volume is exogenous at the 1% level and therefore use IV regression. Further, we apply an overidentification test to test the validity of overidentifying instruments. We cannot reject the null hypothesis ($p > 0.05$) that the instruments are valid for the last two models – our preferred specifications.

The IV coefficients (Table 2, models (5)-(8)) reflect the causal effect of volume on outcome. The coefficients are on average two times higher than the OLS coefficients and still highly significant.

Our preferred specification is model (8) with a ln case volume coefficient of -0.034,[12] indicating that an increase of 1% in case volume reduces the probability of death by 0.034 percentage points (pp). To be more precise: A patient[13] who is treated in a hospital with 70 cases has a probability of death of 7.16%. An increase of 10 cases reduces the probability of death by 0.45 pp to 6.71%.

[12]The expected change in y_{ihc} that is associated with a x% increase in case volume can be calculated as follows: $\widehat{\beta}_1 \cdot ln([100 + p]/100)$.

[13]We take the 'average' patient and set all variables of the model except case volume at their means.

Table 2: Regression results

	LPM				**IV**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln case volume	−0.0081***	−0.0124***	−0.0145***	−0.0141***	−0.0157***	−0.0250***	−0.0327***	−0.0338***
	(0.0016)	(0.0017)	(0.0025)	(0.0025)	(0.0033)	(0.0036)	(0.0054)	(0.0060)
Patient characteristics	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Hospital characteristics	No	No	Yes	Yes	No	No	Yes	Yes
County type indicators	No	No	No	Yes	No	No	No	Yes
R^2	0.000	0.103	0.104	0.104	0.000	0.103	0.103	0.103
First-stage F-statistic					72.515	74.594	49.812	46.790
Test for endogeneity (p-value)					0.005	0.000	0.000	0.000
Overidentification test (p-value)					0.029	0.004	0.084	0.089
Observations	89541	89541	89541	89541	89541	89541	89541	89541
Number of hospitals	1238	1238	1238	1238	1238	1238	1238	1238

Notes: Clustered standard errors (on hospital level) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5 Conclusion

In this paper, we examine the causal relationship of volume on outcome for hip fracture patients. We use a full-sample of all inpatients in Germany from 2007. To overcome endogeneity concerns due to reverse causality and omitted variable bias, we apply an instrumental variable approach, which allows us to assess the plain effect of volume on outcome excluding any effects of selective-referral or unobserved patient heterogeneity. We find evidence for a causal relationship from volume on outcome. Our results indicate that an exemplary increase from 70 to 80 cases decreases the probability of death by 0.45 pp.

With this we are the first who provide evidence for the practice-makes-perfect hypothesis using German data. Analysing the causal relationship of the volume-outcome effect is essential as otherwise policy implications may yield into the wrong direction. We contribute to the debate on minimum volume regulations in Germany by showing that volume is a driving factor for quality and hence, minimum volume regulations would be beneficiary. This evidence is needed as there has been heavy criticism in Germany if the law of minimum volume regulations can really achieve its goal because of these missing confirmations. Based on recent evidence it can also be confirmed that concerns due to endangering access are not at stake in Germany (Hentschker and Mennicken, 2014; Mennicken et al., 2014), so access cannot be the reason for refraining from standards.

Before its introduction or rather further enhancements more analyses are needed for deriving adequate minimum volume standards, i.e. to determine the thresholds at which significant quality differences are observable. What remains unclear are the possible consequences on quality of a higher concentration in the market. Economic theory states that with decreasing competition quality declines. The empirical literature confirms this theory (e.g. Kessler and McClellan, 2000; Gaynor et al., 2013). There are some studies which come to different results (e.g. Gowrisankaran and Town, 2003; Mukamel et al., 2001). But so far, most of the studies show a negative effect of concentration on outcome (Gaynor and Town, 2011). The trade-off between better quality through high volume and lower quality through fewer providers still must be solved. One solution might be starting with low minimum volume standards. Consequently, only hospitals with the lowest volume drop out of the market with negligible consequences for its concentration.

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.

Allareddy, V., Ward, M. M., Wehby, G. L. and Konety, B. R. (2012). The connection between selective referrals for radical cystectomy and radical prostatectomy and volume-outcome effects: an instrumental variables analysis, *American Journal of Medical Quality* **27**(5): 434–440.

Avdic, D., Lundborg, P. and Vikström, J. (2014). Learning-by-doing in a highly skilled profession when stakes are high: Evidence from advanced cancer surgery, *IFAU - Institute for Evaluation of Labour Market and Education Policy* (No. 7).

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.

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 Qualiätssicherung (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., Moreno-Serra, R. and Propper, C. (2013). Death by market power: reform, competition, and patient outcomes in the National Health Service, *American Economic Journal: Economic Policy* **5**(4): 134–166.

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.

Gaynor, M. and Town, R. J. (2011). Chapter nine: competition in health care markets, *in* M. V. Pauly, T. G. Mcguire and P. P. Barros (eds), *Handbook of health economics: Volume 2*, Elsevier, Amsterdam et al., pp. 499–637.

Gowrisankaran, G. and Town, R. J. (2003). Competition, payers, and hospital quality, *Health Services Research* **38**(6p1): 1403–1422.

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. (2014). The volume-outcome relationship and minimum volume standards - Empirical evidence for Germany, *Health Economics* (forthcoming).

Hentschker, C., Mennicken, R. and Schmid, A. (2014). Defining hospital markets: an application to the German hospital sector, *Health Economics Review* **4**: 28.

Iezzoni, L. I. (2003). Reasons for riks adjustment, *in* L. I. Iezzoni (ed.), *Risk adjustment for measuring health care outcomes*, Health Administration Press, Chicago, pp. 1–16.

Keeler, E. B., Rubenstein, L. V., Kahn, K. L., Draper, D., Harrison, E. R., McGinty, M. J., Rogers, W. H. and Brook, R. H. (1992). Hospital characteristics and quality of care, *The Journal of the American Medical Association* **268**(13): 1709–1714.

Kessler, D. P. and McClellan, M. B. (2000). Is hospital competition socially wasteful?, *The Quarterly Journal of Economics* **115**(2): 577–615.

Kuntz, L., Mennicken, R. and Scholtes, S. (2014). Stress on the ward: evidence of safety tipping points in hospitals, *Management Science* (forthcoming).

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.

Mennicken, R., Kolodziej, I. W. K., Augurzky, B. and Kreienberg, R. (2014). Concentration of gynaecology and obstetrics in Germany: Is comprehensive access at stake?, *Health Policy* (forthcoming).

Milcent, C. (2005). Hospital ownership, reimbursement systems and mortality rates, *Health Economics* **14**(11): 1151–1168.

Mukamel, D. B., Zwanziger, J. and Tomaszewski, K. J. (2001). HMO penetration, competition, and risk-adjusted hospital mortality, *Health Services Research* **36**(6 Pt 1): 1019–1035.

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.

Schmid, A. and Ulrich, V. (2013). Consolidation and concentration in the German hospital market: the two sides of the coin, *Health Policy* **109**(3): 301–310.

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.

A Appendix

Table A1: Descriptive statistics of HIP – Elixhauser comorbidities

	Mean	**S.D.**	**Min**	**Max**
Congestive heart failure	0.216	0.411	0	1
Cardiac arrhythmias	0.186	0.389	0	1
Valvular disease	0.043	0.204	0	1
Pulmonary circulation disorders	0.018	0.132	0	1
Peripheral Vascular Disease	0.047	0.212	0	1
Hypertension, uncomplicated	0.451	0.498	0	1
Hypertension, complicated	0.071	0.256	0	1
Paralysis	0.043	0.203	0	1
Other neurological disorders	0.086	0.281	0	1
Chronic pulmonary disease	0.083	0.276	0	1
Diabetes, uncomplicated	0.152	0.359	0	1
Diabetes, complicated	0.049	0.215	0	1
Hypothyroidism	0.053	0.224	0	1
Renal failure	0.128	0.334	0	1
Liver disease	0.017	0.131	0	1
Peptic ulcer disease excluding bleeding	0.003	0.054	0	1
Lymphoma	0.003	0.055	0	1
Metastatic cancer	0.009	0.094	0	1
Solid tumor without metastasis	0.020	0.140	0	1
Rheumatoid arthritis/collagen vascular diseases	0.014	0.119	0	1
Coagulopathy	0.034	0.180	0	1
Obesity	0.054	0.225	0	1
Weight loss	0.025	0.158	0	1
Fluid and electrolyte disorder	0.225	0.418	0	1
Blood loss anemia	0.015	0.123	0	1
Deficiency anemia	0.016	0.127	0	1
Alcohol abuse	0.036	0.185	0	1
Drug abuse	0.015	0.122	0	1
Psychoses	0.010	0.097	0	1
Depression	0.055	0.227	0	1
Observations		89541		

Table A2: Regression results

	LPM				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln case volume	−0.0081***	−0.0124***	−0.0145***	−0.0141***	−0.0157***	−0.0250***	−0.0327***	−0.0338***
	(0.0016)	(0.0017)	(0.0025)	(0.0025)	(0.0033)	(0.0036)	(0.0054)	(0.0060)
Age		0.0023***	0.0023***	0.0023***		0.0023***	0.0023***	0.0023***
		(0.0001)	(0.0001)	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Male		0.0316***	0.0316***	0.0316***		0.0315***	0.0316***	0.0316***
		(0.0022)	(0.0022)	(0.0022)		(0.0022)	(0.0022)	(0.0022)
Admission reason: emergency		0.0009	−0.0002	−0.0001		0.0018	0.0003	0.0002
		(0.0020)	(0.0021)	(0.0021)		(0.0020)	(0.0020)	(0.0020)
transfer		0.0246***	0.0243***	0.0238***		0.0260***	0.0236***	0.0232***
		(0.0080)	(0.0078)	(0.0078)		(0.0076)	(0.0073)	(0.0074)
Femoral neck fracture		0.0020	0.0015	0.0015		0.0013	0.0010	0.0009
		(0.0015)	(0.0015)	(0.0015)		(0.0016)	(0.0016)	(0.0016)
Transfer between departments		0.0256***	0.0290***	0.0294***		0.0276***	0.0291***	0.0295***
		(0.0030)	(0.0030)	(0.0030)		(0.0030)	(0.0029)	(0.0029)
Winter		0.0049***	0.0048***	0.0048***		0.0050***	0.0047***	0.0046***
		(0.0016)	(0.0016)	(0.0016)		(0.0016)	(0.0016)	(0.0016)
Weekend		0.0015	0.0016	0.0016		0.0017	0.0018	0.0018
		(0.0017)	(0.0017)	(0.0017)		(0.0017)	(0.0017)	(0.0017)
Ownership: private not-for-profit			−0.0003	0.0005			−0.0006	−0.0006
			(0.0026)	(0.0026)			(0.0025)	(0.0026)
private for-profit			0.0025	0.0028			0.0003	0.0004
			(0.0034)	(0.0035)			(0.0036)	(0.0037)
University hospital			0.0021	0.0027			0.0006	−0.0006
			(0.0073)	(0.0074)			(0.0069)	(0.0071)
Teaching hospital			0.0036	0.0040			0.0084***	0.0082***
			(0.0026)	(0.0027)			(0.0029)	(0.0028)
Beds: 201-499			0.0094***	0.0090***			0.0170***	0.0166***
			(0.0031)	(0.0031)			(0.0037)	(0.0038)
≥ 500			0.0118***	0.0117***			0.0240***	0.0238***
			(0.0041)	(0.0041)			(0.0051)	(0.0052)
ICU			−0.0124***	−0.0126***			−0.0110***	−0.0114***

Continued on the next page.

	LPM				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			(0.0023)	(0.0023)			(0.0023)	(0.0023)
HHI			−0.0047	−0.0053			0.0057	0.0087
			(0.0052)	(0.0053)			(0.0057)	(0.0063)
Constant	0.0995***	−0.1074***	−0.1005***	−0.1040***	0.1337***	−0.0521***	−0.0357*	−0.0307
	(0.0075)	(0.0098)	(0.0115)	(0.0118)	(0.0151)	(0.0170)	(0.0204)	(0.0239)
Elixhauser comorbidities	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County type indicators	No	No	No	Yes	No	No	No	Yes
R^2	0.000	0.103	0.104	0.104	0.000	0.103	0.103	0.103
First-stage F-statistic					72.515	74.594	49.812	46.790
Test for endogeneity (p-value)					0.005	0.000	0.000	0.000
Overidentification test (p-value)					0.029	0.004	0.084	0.089
Observations	89541	89541	89541	89541	89541	89541	89541	89541
Number of hospitals	1238	1238	1238	1238	1238	1238	1238	1238

Notes: Clustered standard errors (on hospital level) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Regression results with Charlson comorbidity index

	LPM				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln case volume	−0.0081***	−0.0126***	−0.0135***	−0.0133***	−0.0157***	−0.0226***	−0.0279***	−0.0298***
	(0.0016)	(0.0017)	(0.0023)	(0.0024)	(0.0033)	(0.0035)	(0.0053)	(0.0059)
Age		0.0025***	0.0025***	0.0025***		0.0025***	0.0025***	0.0025***
		(0.0001)	(0.0001)	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Male		0.0365***	0.0364***	0.0364***		0.0362***	0.0363***	0.0363***
		(0.0022)	(0.0022)	(0.0022)		(0.0022)	(0.0022)	(0.0022)
Admission reason: emergency		−0.0003	−0.0013	−0.0013		0.0003	−0.0009	−0.0011
		(0.0021)	(0.0021)	(0.0021)		(0.0021)	(0.0021)	(0.0021)
transfer		0.0272***	0.0270***	0.0265***		0.0280***	0.0261***	0.0254***
		(0.0079)	(0.0077)	(0.0078)		(0.0076)	(0.0074)	(0.0074)
Femoral neck fracture		0.0040**	0.0035**	0.0035**		0.0036**	0.0032**	0.0031**
		(0.0016)	(0.0016)	(0.0016)		(0.0016)	(0.0016)	(0.0016)
Transfer between departments		0.0373***	0.0418***	0.0421***		0.0391***	0.0420***	0.0423***
		(0.0033)	(0.0033)	(0.0033)		(0.0032)	(0.0033)	(0.0033)
Winter		0.0055***	0.0054***	0.0054***		0.0055***	0.0053***	0.0052***
		(0.0017)	(0.0017)	(0.0017)		(0.0017)	(0.0017)	(0.0017)
Weekend		0.0011	0.0013	0.0013		0.0011	0.0014	0.0015
		(0.0017)	(0.0017)	(0.0017)		(0.0017)	(0.0017)	(0.0017)
CCI: 1-2		0.0228***	0.0226***	0.0225***		0.0226***	0.0226***	0.0226***
		(0.0015)	(0.0015)	(0.0015)		(0.0015)	(0.0015)	(0.0015)
3-4		0.0664***	0.0660***	0.0659***		0.0661***	0.0659***	0.0658***
		(0.0030)	(0.0030)	(0.0030)		(0.0030)	(0.0030)	(0.0030)
≥ 5		0.1373***	0.1365***	0.1365***		0.1378***	0.1364***	0.1361***
		(0.0057)	(0.0058)	(0.0058)		(0.0057)	(0.0057)	(0.0057)
Ownership: private not-for-profit			−0.0013	−0.0008			−0.0014	−0.0016
			(0.0025)	(0.0025)			(0.0025)	(0.0026)
private for-profit			0.0018	0.0019			0.0003	0.0001
			(0.0035)	(0.0035)			(0.0036)	(0.0037)
University hospital			0.0059	0.0063			0.0057	0.0042
			(0.0075)	(0.0076)			(0.0071)	(0.0073)
Teaching hospital			0.0047*	0.0048*			0.0087***	0.0085***

Continued on the next page.

	LPM				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			(0.0026)	(0.0026)			(0.0028)	(0.0028)
Beds: 201-499			0.0070**	0.0066**			0.0131***	0.0131***
			(0.0031)	(0.0030)			(0.0037)	(0.0038)
≥ 500			0.0087**	0.0083**			0.0181***	0.0184***
			(0.0042)	(0.0042)			(0.0052)	(0.0053)
ICU			−0.0154***	−0.0158***			−0.0143***	−0.0148***
			(0.0024)	(0.0024)			(0.0024)	(0.0024)
HHI			−0.0067	−0.0059			0.0020	0.0059
			(0.0053)	(0.0054)			(0.0059)	(0.0065)
Constant	0.0995***	−0.1324***	−0.1268***	−0.1278***	0.1337***	−0.0884***	−0.0762***	−0.0669***
	(0.0075)	(0.0098)	(0.0112)	(0.0116)	(0.0151)	(0.0164)	(0.0199)	(0.0234)
County type indicators	No	No	No	Yes	No	No	No	Yes
R^2	0.000	0.051	0.052	0.052	0.000	0.050	0.051	0.051
First-stage F-statistic					72.515	73.526	49.606	46.675
Test for endogeneity (p-value)					0.005	0.000	0.002	0.001
Overidentification test (p-value)					0.029	0.005	0.208	0.238
Observations	89541	89541	89541	89541	89541	89541	89541	89541
Number of hospitals	1238	1238	1238	1238	1238	1238	1238	1238

Notes: Clustered standard errors (on hospital level) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Regression results (probit) – marginal effects

	Probit				IV-Probit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln case volume	−0.0081***	−0.0119***	−0.0133***	−0.0131***	−0.0170***	−0.0259***	−0.0315***	−0.0323***
	(0.0016)	(0.0016)	(0.0022)	(0.0022)	(0.0036)	(0.0039)	(0.0056)	(0.0061)
Age		0.0034***	0.0033***	0.0034***		0.0034***	0.0034***	0.0034***
		(0.0001)	(0.0001)	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Male		0.0317***	0.0317***	0.0317***		0.0323***	0.0321***	0.0321***
		(0.0022)	(0.0022)	(0.0022)		(0.0022)	(0.0022)	(0.0022)
Admission reason: emergency		0.0004	−0.0003	−0.0002		0.0019	0.0002	0.0001
		(0.0019)	(0.0020)	(0.0020)		(0.0020)	(0.0020)	(0.0020)
transfer		0.0247***	0.0250***	0.0246***		0.0257***	0.0243***	0.0240***
		(0.0082)	(0.0081)	(0.0081)		(0.0077)	(0.0076)	(0.0077)
Femoral neck fracture		0.0007	0.0003	0.0003		0.0002	−0.0002	−0.0002
		(0.0015)	(0.0015)	(0.0015)		(0.0015)	(0.0015)	(0.0015)
Transfer between departments		0.0249***	0.0302***	0.0306***		0.0276***	0.0306***	0.0310***
		(0.0026)	(0.0028)	(0.0028)		(0.0027)	(0.0027)	(0.0027)
Winter		0.0044***	0.0043***	0.0043***		0.0043***	0.0042***	0.0042***
		(0.0016)	(0.0016)	(0.0016)		(0.0016)	(0.0016)	(0.0016)
Weekend		0.0012	0.0014	0.0013		0.0013	0.0015	0.0015
		(0.0017)	(0.0017)	(0.0017)		(0.0017)	(0.0017)	(0.0017)
Ownership: private not-for-profit			−0.0006	0.0000			−0.0019	−0.0019
			(0.0024)	(0.0024)			(0.0025)	(0.0026)
private for-profit			0.0030	0.0033			0.0005	0.0006
			(0.0034)	(0.0034)			(0.0036)	(0.0037)
University hospital			−0.0002	0.0002			−0.0026	−0.0034
			(0.0072)	(0.0073)			(0.0067)	(0.0068)
Teaching hospital			0.0022	0.0025			0.0066**	0.0063**
			(0.0025)	(0.0025)			(0.0028)	(0.0027)
Beds: 201-499			0.0090***	0.0088***			0.0157***	0.0154***
			(0.0027)	(0.0027)			(0.0031)	(0.0032)
≥ 500			0.0119***	0.0119***			0.0238***	0.0234***
			(0.0038)	(0.0037)			(0.0048)	(0.0048)
ICU			−0.0145***	−0.0147***			−0.0131***	−0.0135***

Continued on the next page.

	Probit				IV-Probit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			(0.0023)	(0.0023)			(0.0023)	(0.0023)
HHI			−0.0048	−0.0054			0.0043	0.0062
			(0.0050)	(0.0051)			(0.0058)	(0.0064)
Elixhauser comorbidities	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County type indicators	No	No	No	Yes	No	No	No	Yes
Observations	89541	89541	89541	89541	89541	89541	89541	89541

Notes: Marginal effects; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: First-stage regression explaining ln case volume

	(1)	(2)	(3)	(4)
Potential patients between 0 to 10 min	0.0042***	0.0042***	0.0028***	0.0026***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
10 to 20 min	0.0008***	0.0008***	0.0006***	0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
20 to 30 min	0.0005***	0.0004***	0.0003***	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Further hospitals between 0 to 15 min	−0.1332***	−0.1335***	−0.0994***	−0.1013***
	(0.0098)	(0.0095)	(0.0081)	(0.0082)
15 to 30 min	−0.0318***	−0.0309***	−0.0233***	−0.0239***
	(0.0047)	(0.0047)	(0.0041)	(0.0042)
Age		−0.0000	0.0008***	0.0007***
		(0.0003)	(0.0002)	(0.0002)
Male		0.0112**	0.0053	0.0044
		(0.0045)	(0.0037)	(0.0036)
Admission reason: emergency		0.0805***	0.0286**	0.0274**
		(0.0192)	(0.0132)	(0.0128)
transfer		0.0428	−0.0207	−0.0155
		(0.0437)	(0.0464)	(0.0434)
Femoral neck fracture		−0.0330***	−0.0263***	−0.0263***
		(0.0048)	(0.0037)	(0.0037)
Transfer between departments		0.1025***	0.0166	0.0170
		(0.0222)	(0.0174)	(0.0165)
Winter		−0.0052	−0.0081**	−0.0083**
		(0.0042)	(0.0034)	(0.0033)
Weekend		0.0070*	0.0083***	0.0085***
		(0.0040)	(0.0032)	(0.0031)
Ownership: private not-for-profit			−0.0506*	−0.0605**
			(0.0280)	(0.0273)
private for-profit			−0.1632***	−0.1630***
			(0.0403)	(0.0407)
University hospital			−0.1194	−0.1536*
			(0.0810)	(0.0838)
Teaching hospital			0.1789***	0.1610***
			(0.0293)	(0.0294)
Beds: 201-499			0.3919***	0.3883***
			(0.0377)	(0.0373)
≥ 500			0.6238***	0.6059***
			(0.0515)	(0.0500)
ICU			0.0061	−0.0029
			(0.0308)	(0.0301)
HHI			0.4629***	0.4818***
			(0.0569)	(0.0569)
Constant	4.2356***	4.1720***	3.5898***	3.6261***
	(0.0245)	(0.0334)	(0.0448)	(0.0738)
Elixhauser comorbidities	No	Yes	Yes	Yes
County type indicators	No	No	No	Yes
R^2	0.271	0.283	0.524	0.535
First-stage F-statistic	72.515	74.594	49.812	46.790
Test for endogeneity (p-value)	0.005	0.000	0.000	0.000
Overidentification test (p-value)	0.029	0.004	0.084	0.089
Observations	89541	89541	89541	89541
Number of hospitals	1238	1238	1238	1238

Notes: Clustered standard errors (on hospital level) in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.