Separate & Concentrate: Accounting for Patient Complexity in General Hospitals

Supplementary Material

Ludwig Kuntz
Faculty of Management, Economics and Social Science, University of Cologne, kuntz@wiso.uni-koeln.de
Stefan Scholtes
Judge Business School, University of Cambridge, s.scholtes@jbs.cam.ac.uk
Sandra Sülz
Erasmus School of Health Policy and Management, Erasmus University Rotterdam, sulz@eshpm.eur.nl

---

This supplementary material accompanies the paper Separate & Concentrate: Accounting for Patient Complexity in General Hospitals. Its main purpose is to provide additional explanation and present results of different model specifications, variable definitions and sample exclusion criteria. This report complements the main paper and should not be read in isolation.

Contents

1. Disease segments

Table 1: 39 disease segments used in the main paper (ICD blocks with mortality $\geq 1\%$)

Description of ICD Block	% sample patients	Mortality	# different hospitals
Certain infectious and parasitic diseases			
Other bacterial diseases	1.719%	6.408%	60
Protozoal diseases	0.009%	4.348%	7
Other infectious diseases	0.210%	1.434%	46
Neoplasms			
Malignant neoplasms of lip, oral cavity and pharynx	1.225%	1.170%	39
Malignant neoplasms of digestive organs	7.813%	2.549%	60
Malignant neoplasms of respiratory and intrathoracic organs	5.988%	3.445%	60
Malignant neoplasms of bone and articular cartilage	0.101%	1.493%	22
Malignant neoplasms of mesothelial and soft tissue	0.500%	1.659%	48
Malignant neoplasm of breast	3.725%	1.023%	59
Malignant neoplasms of male genital organs	2.999%	1.094%	56
Malignant neoplasms of eye, brain and other parts of central nervous system	0.841%	1.615%	50
Malignant neoplasms of ill-defined, secondary and unspecified sites	5.001%	2.285%	60
Malignant neoplasms, stated or presumed to be primary, of lymphoid, hematopoietic and related tissue	3.859%	1.290%	59
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism			
Aplastic and other anaemias	0.407%	1.298%	59
Coagulation defects, purpura and other hemorrhagic conditions	0.396%	1.621%	58
Other diseases of blood and blood-forming organs	0.172%	1.316%	47
Endocrine, nutritional and metabolic diseases			
Metabolic disorders	1.177%	3.462%	60
Mental and behavioral disorders			
Organic, including symptomatic, mental disorders	0.719%	1.312%	57
Diseases of the nervous system			
Systemic atrophies primarily affecting the central nervous system	0.127%	5.935%	28
Other disorders of the nervous system	0.339%	1.891%	46
Diseases of the circulatory system			
Ischemic heart diseases	12.649%	2.174%	60
Pulmonary heart disease and diseases of pulmonary circulation	0.867%	9.352%	59
Other forms of heart disease	11.994%	3.333%	60
Cerebrovascular diseases	7.291%	4.806%	59

Description of ICD Block	% sample patients	Mortality	# different hospitals
Diseases of arteries, arterioles and capillaries	4.228%	1.891%	58
Diseases of the respiratory system			
Influenza and pneumonia	3.943%	7.280%	60
Chronic lower respiratory diseases	3.364%	1.985%	60
Lung diseases due to external agents	0.400%	19.906%	59
Other respiratory diseases principally affecting the interstitium	0.317%	1.070%	44
Suppurative and necrotic conditions of lower respiratory tract	0.142%	1.596%	41
Other diseases of pleura	0.556%	1.494%	59
Other diseases of the respiratory system	0.812%	5.481%	56
Diseases of the digestive system			
Other diseases of intestines	6.404%	2.150%	60
Diseases of peritoneum	0.527%	2.434%	56
Diseases of liver	1.550%	4.964%	59
Other diseases of the digestive system	1.142%	3.301%	60
Diseases of the genitourinary system			
Renal failure	1.330%	4.962%	59
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified			
General symptoms and signs	3.315%	1.900%	60
Injury, poisoning and certain other consequences of external causes			
Injuries to the hip and thigh	2.904%	1.727%	57

2. Model covariates

Table 2: **Summary statistics and correlations of covariates**

Variable	Mean	Death	High volume	High focus	High con.
Independent variables of interest					
High volume	0.8604	−0.0516∗			
High focus	0.7258	−0.0296∗	0.2217∗		
High concentration	0.3113	0.0287∗	−0.2620∗	0.0602∗	
Routine patient	0.4194	−0.0699∗	0.0940∗	0.1583∗	0.0021
Complex patient	0.1375	0.0594∗	−0.0477∗	−0.1096∗	−0.0162∗
Benchmark patient	0.4431	0.0283∗	−0.0603∗	−0.0813∗	0.0091∗
Patient health status					
Age	66.5206	0.1078∗	−0.1150∗	−0.0589∗	0.0704∗
Age squared	4641.4770	0.1177∗	−0.1310∗	−0.0759∗	0.0776∗
Male	0.5396	−0.0184∗	0.0599∗	0.0406∗	−0.0485∗
Congestive heart failure	0.1432	0.0520∗	−0.0715∗	−0.0904∗	−0.0072∗
Cardiac arrhythmias	0.1774	0.0366∗	−0.0560∗	−0.0971∗	−0.0002
Valvular disease	0.0701	−0.0147∗	−0.0065∗	−0.0765∗	−0.0422∗
Pulmonary circulation disorders	0.0251	0.0174∗	−0.0157∗	−0.0546∗	−0.0210∗
Peripheral vascular disorders	0.0752	0.0002	0.0352∗	0.0284∗	−0.0452∗
Hypertension, uncomplicated	0.3870	−0.0511∗	−0.0121∗	−0.0130∗	−0.0255∗
Hypertension, complicated	0.0667	−0.0151∗	0.0070∗	−0.0314∗	−0.0362∗
Paralysis	0.0556	0.0366∗	−0.0008	−0.0004	−0.0448∗
Other neurological disorders	0.0555	0.0260∗	−0.0286∗	−0.0172∗	−0.0156∗
Chronic pulmonary disease	0.1053	0.0056∗	−0.0054∗	−0.0178∗	−0.0125∗
Diabetes, uncomplicated	0.1284	−0.0010	−0.0199∗	−0.0223∗	−0.0039∗
Diabetes, complicated	0.0678	0.0160∗	−0.0397∗	−0.0152∗	0.0186∗
Hypothyroidism	0.0321	−0.0171∗	0.0054∗	−0.0200∗	−0.0248∗
Renal failure	0.0982	0.0334∗	−0.0189∗	−0.0517∗	−0.0251∗
Liver disease	0.0368	0.0158∗	0.0015	−0.0110∗	0.0015
Peptic ulcer disease excluding bleeding	0.0046	−0.0037	−0.0027	−0.0053∗	−0.0021
AIDS/HIV	0.0008	−0.0025	0.0097∗	−0.0053∗	−0.0106∗
Lymphoma	0.0076	0.0005	0.0081∗	0.0032	−0.0184∗
Metastatic cancer	0.1250	0.0243∗	0.0871∗	0.1428∗	0.0281∗
Solid tumor without metastasis	0.0641	0.0003	0.0426∗	0.0555∗	−0.0346∗
Rheumatoid arthritis/collagen, vascular diseases	0.0116	−0.0086∗	−0.0008	−0.0153∗	−0.0087∗
Coagulopathy	0.0322	0.0183∗	0.0325∗	0.0057∗	−0.0435∗
Obesity	0.0871	−0.0327∗	−0.0050∗	−0.0019	−0.0213∗
Weight loss	0.0311	0.0676∗	−0.0366∗	0.0059∗	0.0166∗
Fluid and electrolyte disorders	0.1260	0.0645∗	−0.0662∗	−0.0532∗	−0.0050∗
Blood loss anemia	0.0056	−0.0003	−0.0191∗	−0.0236∗	0.0129∗
Deficiency anemias	0.0133	−0.0023	−0.0161∗	−0.0210∗	0.0097∗
Alcohol abuse	0.0324	0.0034	−0.0095∗	−0.0169∗	−0.0015
Drug abuse	0.0040	−0.0079∗	−0.0011	−0.0082∗	−0.0072∗
Psychoses	0.0044	0.0003	−0.0124∗	−0.0137∗	−0.0012
Depression	0.0324	−0.0177∗	−0.0239∗	−0.0082∗	0.0007

Variable	Mean	Death	High volume	High focus	High con.
Socioeconomic factors for patient's region					
Employment rate	74.4727	0.0164∗	−0.0792∗	0.0544∗	0.1883∗
Residents per physician	666.7671	0.0047∗	−0.1099∗	0.0819∗	0.1480∗
Population density	1092.5460	−0.0143∗	0.1466∗	−0.0671∗	−0.1357∗
GDP per capita	25.4017	0.0195∗	0.0698∗	−0.0564∗	0.0184∗
Admission day of the week					
Sunday	0.0720	0.0365∗	−0.0361∗	−0.0424∗	0.0090∗
Monday	0.2124	−0.0204∗	0.0186∗	0.0279∗	0.0019
Tuesday	0.1947	−0.0201∗	0.0257∗	0.0302∗	−0.0028
Wednesday	0.1790	−0.0126∗	0.0186∗	0.0223∗	−0.0026
Thursday	0.1603	−0.0090∗	0.0110∗	0.0080∗	−0.0094∗
Friday	0.1220	0.0169∗	−0.0200∗	−0.0242∗	−0.0041∗
Saturday	0.0596	0.0399∗	−0.0552∗	−0.0674∗	0.0161∗
Admission, month-year					
1-2004	0.0301	0.0027	−0.0239∗	0.0019	0.0244∗
2-2004	0.0267	0.0045∗	−0.0197∗	−0.0018	0.0234∗
3-2004	0.0312	0.0049∗	−0.0202∗	−0.0036	0.0220∗
4-2004	0.0269	0.0040∗	−0.0204∗	0.0001	0.0252∗
5-2004	0.0262	0.0012	−0.0193∗	−0.0027	0.0245∗
6-2004	0.0282	0.0012	−0.0172∗	0.0005	0.0245∗
7-2004	0.0276	0.0018	−0.0198∗	0.0013	0.0302∗
8-2004	0.0262	0.0058∗	−0.0193∗	−0.0017	0.0315∗
9-2004	0.0270	0.0044∗	−0.0191∗	0.0036	0.0277∗
10-2004	0.0267	0.0024	−0.0216∗	−0.0010	0.0257∗
11-2004	0.0289	0.0052∗	−0.0160∗	0.0045∗	0.0269∗
12-2004	0.0143	0.0012	−0.0236∗	−0.0012	0.0201∗
1-2005	0.0666	−0.0025	0.0140∗	0.0011	−0.0165∗
2-2005	0.0613	0.0015	0.0024	−0.0066∗	−0.0125∗
3-2005	0.0662	0.0069∗	0.0004	−0.0042∗	−0.0121∗
4-2005	0.0628	−0.0019	0.0208∗	0.0040∗	−0.0209∗
5-2005	0.0617	−0.0023	0.0140∗	−0.0000	−0.0167∗
6-2005	0.0612	−0.0069∗	0.0182∗	0.0002	−0.0241∗
7-2005	0.0598	−0.0056∗	0.0151∗	0.0053∗	−0.0186∗
8-2005	0.0609	−0.0037	0.0234∗	0.0032	−0.0252∗
9-2005	0.0575	−0.0049∗	0.0180∗	−0.0037	−0.0230∗
10-2005	0.0592	−0.0019	0.0185∗	0.0013	−0.0186∗
11-2005	0.0607	−0.0054∗	0.0153∗	−0.0015	−0.0172∗
12-2005	0.0021	−0.0012	0.0009	0.0034	−0.0006
Federal state					
1	0.0068	0.0010	0.0221∗	0.0218∗	0.0004
2	0.0474	0.0070∗	−0.1417∗	0.0108∗	0.1212∗
3	0.1032	−0.0305∗	0.0926∗	−0.0476∗	−0.1547∗
4	0.0839	−0.0163∗	0.0834∗	−0.0231∗	−0.1185∗
5	0.0403	0.0044∗	0.0774∗	−0.0023	−0.0955∗
6	0.0192	0.0104∗	−0.0414∗	−0.0274∗	0.0377∗
7	0.0442	0.0017	0.0623∗	−0.0176∗	−0.0282∗
8	0.0159	−0.0009	0.0110∗	0.0177∗	0.0346∗

Variable	**Mean**	**Death**	**High volume**	**High focus**	**High con.**
9	0.0247	−0.0028	−0.0196∗	0.0331∗	−0.0460∗
10	0.2008	0.0220∗	−0.0484∗	−0.0492∗	0.0638∗
11	0.0608	0.0171∗	−0.0786∗	0.0415∗	0.0368∗
12	0.0211	0.0098∗	−0.2500∗	0.0089∗	0.0888∗
13	0.1093	−0.0122∗	−0.0202∗	0.0379∗	0.1206∗
14	0.0597	0.0335∗	−0.0339∗	0.0194∗	0.1665∗
15	0.1069	−0.0107∗	0.0820∗	0.0280∗	−0.0849∗
16	0.0555	−0.0219∗	0.0849∗	0.0032	−0.0946∗

Notes. * p<0.05. Summary and correlation statistics shown at the patient-level.
Segment and Hospital FE not shown.

3. Calculation of dichotomous independent variables

This section explains the details of the dichotomization of the independent variables segment volume, segment focus, and segment concentration.

The volume measure.

We define an indicator variable $X_{ish} = 1$ if patient $i \in \{1, \ldots, N\}$ belongs to segment $s \in \{1, \ldots, S\}$ and is admitted to hospital $h \in \{1, \ldots, H\}$, and 0 otherwise, and calculate the annualized volume of patients in segment s and hospital h as $\text{Vol}_{sh} = \sum_{i=1}^{N} X_{ish}$, for hospitals h with one year of data and as $\text{Vol}_{sh} = \frac{1}{2}\sum_{i=1}^{N} X_{ish}$ for hospitals h with two years of data. The variable Vol_{sh} is then dichotomized at the segment level via a median split across hospitals for each segment s

$$\bar{V}_{sh} = \begin{cases} 1 \text{ if } \text{Vol}_{sh} \geq \text{median}\{\text{Vol}_{sh'} \mid h' \in \{1, \ldots, H\}, \text{Vol}_{sh'} > 0\} \\ 0 \text{ otherwise.} \end{cases} \tag{1}$$

This binary variable $\bar{V}_{sh}$ splits the hospitals into high- and low-volume hospitals for the fixed disease segment s. We assign to patient i in segment $s = s(i)$ and hospital $h = h(i)$ the dichotomous volume variable $V_{sh} = \bar{V}_{s(i)h(i)}$.

The focus measure.

We calculate a continuous focus measure for segment s in hospital h as the annualized relative volume of patients in that segment and hospital:

$$\text{Foc}_{sh} = \frac{\sum_{i=1}^{N} X_{ish}}{\sum_{s' \in \mathcal{S}} \sum_{i=1}^{N} X_{is'h}}, \tag{2}$$

where $X_{ish} = 1$ if patient i is in segment s and admitted to hospital h, and 0 otherwise, and N is the total number of patient episodes in the data. The variable Foc_{sh} is then dichotomized at the segment level via a median split across hospitals for each segment s

$$\bar{F}_{sh} = \begin{cases} 1 \text{ if } \text{Foc}_{sh} \geq \text{median}\{\text{Foc}_{sh'} \mid h' \in \{1, \ldots, H\}, \text{Foc}_{sh'} > 0\} \\ 0 \text{ otherwise.} \end{cases} \tag{3}$$

This binary variable $\bar{F}_{sh}$ splits the hospitals into high- and low-volume hospitals for the fixed disease segment s. We assign to patient i in segment $s = s(i)$ and hospital $h = h(i)$ the dichotomous volume variable $F_{sh} = \bar{F}_{s(i)h(i)}$.

The concentration measure.

Let $\mathcal{D} = \{1, \ldots, D\}$ be the set of all hospital departments in the sample. Denote by $h(d)$ department d's hospital h, by $\mathcal{D}_h \subset \mathcal{D}$ the set of all departments of hospital h, and define the variable $X'_{isd} = 1$ if patient i belongs to segment s and is admitted to hospital department $d \in \mathcal{D}$ and 0 otherwise. Department d's proportion of segment s patients in its hospital $h(d)$ is then defined as

$$\text{Q}_{sd} = \frac{\sum_{i=1}^{N} X'_{isd}}{\sum_{d' \in \mathcal{D}_{h(d)}} \sum_{i=1}^{N} X'_{isd'}}. \tag{4}$$

The hospital h's default department for segment s is defined as the department $d \in \mathcal{D}_h$ that maximizes Q_{sd}, which leads to a continuous measure of departmental concentration for segment s in hospital h as

$$\text{Con}_{sh} = \max\{Q_{sd} \mid d \in \mathcal{D}_{h(d)}\}. \tag{5}$$

We dichotomize this continuous measure by splitting the hospitals that admit patients in segment s into a high- and a low-concentration group for the segment:

$$\bar{C}_{sh} = \begin{cases} 1 \text{ if } \text{Con}_{sh} \geq \text{median}\{\text{Con}_{sh'} \mid h' \in \{1,\ldots,H\}, \sum_{d' \in \mathcal{D}_{h'(d)}} \sum_{i=1}^{N} X'_{isd'} > 0\} \\ 0 \text{ otherwise.} \end{cases} \tag{6}$$

Finally, we assign to a patient i in segment $s = s(i)$ and hospital $h = h(i)$ the dichotomous concentration variable $C_{sh} = \bar{C}_{s(i)h(i)}$.

4. The need to dichotomize

In the main paper we argue that specifying a linear relationship between the patient's latent health status and the volume z-scores is inappropriate because the medical literature provides robust evidence of non-linear relationships between volume and mortality. Specifically, there is evidence that suggests that volume has a beneficial effect on quality up to a volume threshold, beyond which there is no further beneficial effect. We now present an illustrative example to support our specification choice showing that dichotomization is indeed preferable to a misspecified linear model if the underlying relationship is non-linear. To do so, we use two fictitious segments, simulate a skewed volume distribution and calculate the corresponding volume z-scores. For both segments, we then model a non-linear relationship between the volume z-scores and the patient's mortality risk MR, using segment-specific thresholds and slopes, and adding for each observation a simulated error term (confer Figure 1). We then estimate this relationship in two ways, (i) with a linear specification

$$MR = \beta_0 + \beta_S S + \beta_V Volume_{Z-Score} + \epsilon$$

, where S denotes the fixed effect for the segment, and (ii) with dichotomous median-split specification:

$$MR = \beta_0 + \beta_S S + \beta_D Volume_{Dummy} + \epsilon$$

, where $Volume_{Dummy}$ is a binary variable that is equal to 1 if the z-score is higher than the segment-specific median and 0 otherwise.

Figure 2 shows the predictions from these models, separated by segment. The predictions obtained from the median-split specifications are more closely aligned with the underlying

relationship thant the prediction obtained from the linear specification; R^2 is higher for the median-split model ($R^2 = 0.79$) than for the linear model ($R^2 = 0.64$). This lends support to the argument that a median-split model may well provide a better model fit than a linear specification provided the underlying relationship is non-linear.

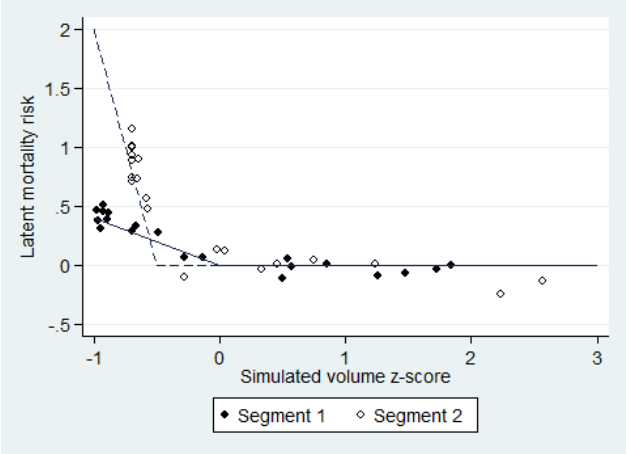
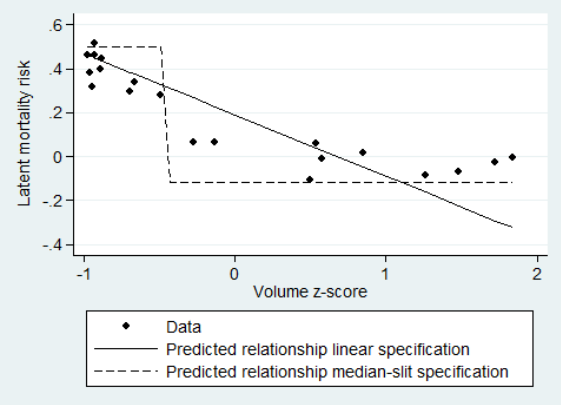
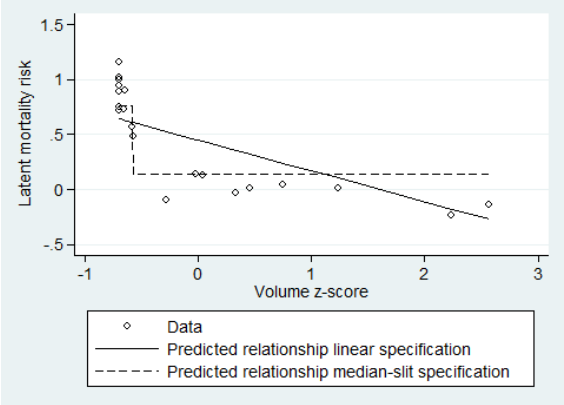


Figure 1 **Simulated relationship between volume z-score and patient's health status for two segments**



(a) Segment 1



(b) Segment 2

Figure 2 **Predictions**

5. Calculation of instrumental variables

In this section, we explain how the instrumental variables (IV) were calculated. We outline the computation for the endogenous volume variable only as the computations were analogous for the focus and concentration variables.

The first IV is a continuous differential distance (DD) variable, defined as the difference between a patient's distance to the nearest high-volume hospital for her disease segment s and the patient's

distance to the nearest hospital that treats patients in segment s, independently of the segment volume in the hospital (McClellan et al. 1994). The variable is zero if the nearest hospital is a high-volume hospital; otherwise it is a measure of the inconvenience for the patient to choose a high-volume hospital over her nearest hospital. While our main database provides information on the patient's place of residence, this data covers only a sample of hospitals and not the total hospital population in Germany. In order to enlarge the sample of hospitals, we therefore use information in the Hospital Quality Reports (Gemeinsamer Bundesausschuss 2016), which are available for 82% of the German general hospitals, to approximate the volume of hospital patients in each segment s. Specifically, the reports list for each clinical departments at least the ten highest volume ICD 3-digit codes. Aggregating the volumes of the listed ICD codes up to the level of ICD blocks (our disease segments), we obtain a lower bound of the volume of hospital patients in each segment and hospital department. In line with our prior dichotomization, we then split all hospitals that treat a fixed segment s into two groups, the 50% of hospitals with the highest approximated volumes in the segment and the remaining hospitals. We assume that, in addition to the hospitals that record patients in segment s in their Quality Report, all hospitals that are classified as general hospitals with a regional service mandate (bed size above 250) in the German hospital plan treat all segments and consider them a low-volume hospital for segment s if the segment is not listed in their Quality Report. Having identified high-volume and low-volume hospitals for each segment, we can define for each patient the nearest high-volume hospital and the nearest permissible hospital based on zip code information of the patient and hospital using the STATA module geonear (Picard 2012). Based upon this information, we generate our first instrumental variable, i.e. the additional distance a patient would have to travel to reach a high-volume hospital for her segment s (McClellan et al. 1994).

We cannot compute exact segment volumes in each hospital and therefore the approximate volume variable and its dichotomization is not identical to the original volume variable in our sample hospitals. However, while this reduces the statistical power of the instrument, the strong conformity (83%) for the two dichotomous variables of the sample hospitals suggests that the instrument will remain valid. Note that the validity of the exclusion restriction is not affected by the approximation. We further alleviate any residual concerns about weak instrumentation by including as a second IV a set of binary variables, based on the idea that a patient's propensity to be admitted to a high-volume hospital is the higher, the more high-volume hospitals there are in the vicinity of her place of residence (KC and Terwiesch 2011). We operationalize this idea with a set of K binary variables D_{ik} ($k \in \{1, \ldots, K\}$) which are equal to 1 if the k-th nearest hospital to patient i is a high-volume hospital for patient i's segment and 0 otherwise. The results are robust

with respect to the number of binary variables K that comprise the instrument and our results are based on $K = 5$.

The critical assumption for the validity of IVs is the exclusion restriction, which requires that a valid IV is uncorrelated with *unobserved* mortality risk factors. While this assumption cannot be tested statistically, Tables 3-5 provide some evidence for the validity of the exclusion restriction for the differential distance by demonstrating that *observed* mortality risk factors, such as age and the presence of the Elixhauser comorbidities, do not differ substantially across different values of differential distance for volume (focus, concentration). We present the results in the tables using a cut off value of 10km due to peculiarities of the German Context. In Germany, distances are generally much lower than, for example, in the USA, and 97.5% of the population live within 20 minutes drive of their nearest general hospital. In our sample, the differential distance (which is the distance between hospitals) is above 10 km only for about 10-20% of the sample, depending on the variable. We therefore found this to be a sensible cut off. We experimented with other cut offs and the results are similar.

Table 6 further outlines how the instrumental variables are related to each other and supports the view that differential distance and the dummy IVs are complementary. Nevertheless, to alleviate concerns about the correlations of the instruments in the three selection equations (i.e. that an instrument in one selection could be an omitted variable in another selection equation and thereby cause bias), we have re-estimated the models with all instruments in all three selection equations; the results in Table 7 confirm the results of the main model. In addition, to alleviate concerns about overfitting, we varied the number of binary variables K that form the set of binary variables D_k and we present the results for $K = 3$ (Table 8), $K = 4$ (Table 9). We also re-estimated a model with a parsimonious IV structure, using only one IV per selection equation ($D1_V, D1_F, D1_C$, respectively) (Table 10). Again, the results are in line with the the results in the main model.

Table 3 Patient characteristics by differential distance to next high-volume hospital

	Differential distance		
	DD=0 km	DD ∈ (0 km; 10 km)	DD≥ 10 km
Mean (SD) age	66.6 (14.7)	66.4 (14.9)	66.2 (14.5)
Male	54%	54%	55%
Routine patients	41%	42%	48%
Benchmark patients	44%	45%	43%
Complex patients	15%	13%	10%
Hypertension, uncomplicated	39%	38%	37%
Cardiac arrhythmias	18%	18%	16%
Congestive heart failure	14%	15%	13%
Diabetes, uncomplicated	13%	13%	12%
Fluid and electrolyte disorders	12%	13%	12%
Metastatic cancer	12%	12%	15%
Chronic pulmonary disease	10%	11%	11%
Renal failure	10%	10%	9%
Obesity	9%	9%	9%
Peripheral vascular disorders	8%	8%	7%
Valvular disease	7%	7%	6%
Solid tumor without metastasis	6%	6%	7%
Diabetes, complicated	6%	8%	8%
Hypertension, complicated	6%	8%	7%
Paralysis	6%	5%	5%
Other neurological disorders	6%	6%	5%
Liver disease	4%	3%	4%
Coagulopathy	3%	3%	3%
Hypothyroidism	3%	3%	3%
Alcohol abuse	3%	3%	3%
Depression	3%	3%	3%
Weight loss	3%	3%	4%
Pulmonary circulation disorders	2%	3%	3%
Deficiency anemias	1%	1%	1%
Rheumatoid arthritis/collagen, vascular diseases	1%	1%	1%
Lymphoma	1%	1%	1%
Blood loss anemia	1%	1%	1%
Peptic ulcer disease excluding bleeding	0%	0%	0%
Psychoses	0%	0%	0%
Drug abuse	0%	0%	0%
AIDS/HIV	0%	0%	0%
high-volume treatment	91%	80%	68%
seven-day mortality	2.98%	3.02%	2.97%
Total patients	178,107	59,613	27,413

Table 4 Patient characteristics by differential distance to next high-focus hospital

	Differential distance		
	DD=0 km	DD ∈ (0 km; 10 km)	DD≥ 10 km
Mean (SD) age	66.6 (14.7)	66.3 (14.8)	66.3 (14.7)
Male	54%	55%	54%
Routine patients	41%	44%	43%
Benchmark patients	45%	44%	43%
Complex patients	15%	12%	13%
Hypertension, uncomplicated	39%	38%	37%
Cardiac arrhythmias	18%	18%	18%
Congestive heart failure	14%	15%	14%
Diabetes, uncomplicated	13%	13%	12%
Fluid and electrolyte disorders	13%	13%	12%
Metastatic cancer	12%	13%	13%
Chronic pulmonary disease	11%	11%	11%
Renal failure	10%	10%	10%
Obesity	9%	9%	9%
Peripheral vascular disorders	8%	7%	8%
Valvular disease	7%	7%	7%
Diabetes, complicated	6%	7%	8%
Hypertension, complicated	6%	7%	8%
Solid tumor without metastasis	6%	6%	8%
Paralysis	6%	5%	6%
Other neurological disorders	6%	5%	6%
Liver disease	4%	3%	4%
Hypothyroidism	3%	3%	3%
Coagulopathy	3%	3%	3%
Depression	3%	3%	3%
Alcohol abuse	3%	3%	4%
Weight loss	3%	3%	3%
Pulmonary circulation disorders	2%	2%	3%
Deficiency anemias	1%	1%	1%
Rheumatoid arthritis/collagen, vascular diseases	1%	1%	1%
Lymphoma	1%	1%	1%
Blood loss anemia	1%	1%	1%
Psychoses	0%	0%	0%
Peptic ulcer disease excluding bleeding	0%	0%	1%
Drug abuse	0%	0%	0%
AIDS/HIV	0%	0%	0%
high-focus treatment	75%	73%	59%
seven-day mortality	2.96%	3.06%	3.01%
Total patients	173,948	59,112	32,073

Table 5 Patient characteristics by differential distance to next high-concentration hospital

	Differential distance		
	DD=0 km	DD ∈ (0 km; 10 km)	DD≥ 10 km
Mean (SD) age	66.5 (14.9)	66.7 (14.6)	66.4 (14.3)
Male	54%	54%	54%
Routine patients	42%	42%	41%
Benchmark patients	44%	45%	45%
Complex patients	14%	13%	14%
Congestive heart failure	39%	39%	38%
Cardiac arrhythmias	17%	18%	18%
Valvular disease	14%	15%	14%
Pulmonary circulation disorders	13%	13%	13%
Peripheral vascular disorders	12%	13%	12%
Hypertension, uncomplicated	12%	13%	13%
Hypertension, complicated	10%	11%	11%
Paralysis	10%	10%	10%
Other neurological disorders	8%	9%	10%
Chronic pulmonary disease	7%	8%	8%
Diabetes, uncomplicated	7%	7%	7%
Diabetes, complicated	7%	7%	6%
Hypothyroidism	6%	7%	8%
Renal failure	6%	7%	9%
Liver disease	5%	6%	5%
Peptic ulcer disease excluding bleeding	5%	6%	6%
AIDS/HIV	4%	3%	4%
Lymphoma	3%	3%	3%
Metastatic cancer	3%	3%	3%
Solid tumor without metastasis	3%	3%	3%
Rheumatoid arthritis/collagen, vascular diseases	3%	4%	3%
Coagulopathy	3%	3%	4%
Obesity	2%	3%	3%
Weight loss	1%	1%	1%
Fluid and electrolyte disorders	1%	1%	1%
Blood loss anemia	1%	1%	1%
Deficiency anemias	1%	1%	1%
Alcohol abuse	0%	0%	0%
Drug abuse	0%	0%	0%
Psychoses	0%	0%	0%
Depression	0%	0%	0%
high-concentration treatment	33%	33%	23%
seven-day mortality	2.96%	3.07%	2.95%
Total patients	146,309	65,572	53,252

Table 6 Correlation of instrumental variables

	DD_V	$D1_V$	$D2_V$	$D3_V$	$D4_V$	$D5_V$	DD_F	$D1_F$	$D2_F$	$D3_F$	$D4_F$	$D5_F$	DD_C	$D1_C$	$D2_C$	$D3_C$	$D4_C$
$D1_V$	-0.51*																
$D2_V$	-0.18*	-0.09*															
$D3_V$	-0.08*	0.07*	0.03*														
$D4_V$	0.02*	0.03*	0.02*	0.03*													
$D5_V$	-0.06*	0.02*	0.06*	0.11*	0.10*												
DD_F	0.54*	-0.25*	-0.12*	-0.07*	-0.00	-0.02*											
$D1_F$	-0.22*	0.48*	-0.07*	-0.01*	0.03*	-0.06*	-0.51*										
$D2_F$	-0.05*	-0.05*	0.52*	0.01*	-0.01*	-0.02*	-0.16*	0.02*									
$D3_F$	-0.01*	0.02*	0.04*	0.56*	0.02*	0.06*	-0.08*	0.04*	0.09*								
$D4_F$	0.07*	-0.02*	-0.02*	0.00	0.62*	0.07*	-0.01*	0.05*	-0.04*	0.01*							
$D5_F$	-0.00	0.02*	0.02*	0.07*	0.09*	0.54*	-0.04*	0.02*	0.00*	0.05*	0.12*						
DD_C	0.08*	0.14*	-0.05*	0.02*	0.04*	0.03*	0.24*	0.03*	0.01*	0.02*	0.03*	-0.02*					
$D1_C$	0.10*	-0.19*	-0.01*	-0.09*	-0.04*	-0.07*	-0.03*	0.08*	0.01*	-0.05*	0.01*	-0.01*	-0.53*				
$D2_C$	0.08*	0.08*	-0.22*	0.03*	-0.03*	-0.03*	0.05*	0.01*	-0.01*	0.03*	-0.03*	0.00*	-0.18*	0.02*			
$D3_C$	0.03*	0.04*	0.02*	-0.15*	-0.02*	-0.05*	0.002	0.05*	0.02*	0.05*	-0.02*	0.03*	-0.11*	-0.00	0.05*		
$D4_C$	0.04*	0.01*	-0.04*	-0.01*	-0.17*	0.01*	0.04*	-0.02*	0.03*	0.03*	-0.03*	0.05*	-0.05*	0.01*	0.05*	0.07*	
$D5_C$	0.04*	-0.02*	-0.01*	-0.06*	0.00	-0.25*	0.01*	0.05*	0.08*	-0.02*	-0.00	0.01*	-0.06*	0.01*	0.03*	0.10*	0.03*

* p<0.05

Table 7 Simultaneous equations models for seven-day mortality: All IVs

Mortality equation	Simultaneous equations model		
Vol	0.037		
	(0.056)		
Vol * PR	0.016		
	(0.041)		
Vol * PC	0.127**		
	(0.041)		
Foc	−0.011		
	(0.051)		
Foc * PR	−0.189***		
	(0.037)		
Foc * PC	0.038		
	(0.033)		
Con	−0.160*		
	(0.066)		
Con * PR	0.089*		
	(0.039)		
Con * PC	−0.069*		
	(0.034)		
Selection equations (Ivs)		Vol	Foc	Con
DD_V		−0.028***	−0.004	0.014+
		(0.006)	(0.007)	(0.007)
$D1_V$		0.435***	−0.117	0.229*
		(0.097)	(0.115)	(0.102)
$D2_V$		0.090	−0.189+	−0.153+
		(0.082)	(0.097)	(0.083)
$D3_V$		0.165*	0.057	0.136
		(0.072)	(0.086)	(0.083)
$D4_V$		−0.031	0.024	0.108
		(0.063)	(0.078)	(0.079)
$D5_V$		0.019	−0.059	−0.061
		(0.062)	(0.081)	(0.073)
DD_F		−0.006	−0.008	−0.009
		(0.007)	(0.008)	(0.009)
$D1_F$		−0.220*	0.538***	−0.194+
		(0.104)	(0.115)	(0.111)
$D2_F$		0.029	0.304***	0.120
		(0.075)	(0.090)	(0.077)
$D3_F$		−0.094	0.087	−0.034
		(0.071)	(0.087)	(0.087)
$D4_F$		−0.085	−0.086	−0.004
		(0.063)	(0.074)	(0.067)
$D5_F$		−0.011	−0.002	0.057
		(0.068)	(0.085)	(0.077)
DD_C		0.022***	0.016**	−0.004
		(0.006)	(0.006)	(0.006)
$D1_C$		0.260**	0.314***	0.350***
		(0.083)	(0.095)	(0.088)
$D2_C$		0.073	0.174*	0.092
		(0.064)	(0.076)	(0.066)
$D3_C$		0.063	0.015	0.169**
		(0.063)	(0.072)	(0.062)
$D4_C$		0.096+	0.138+	0.152*
		(0.058)	(0.074)	(0.067)
$D5_C$		0.058	−0.006	−0.043
		(0.057)	(0.066)	(0.061)
Error correlations				
$\rho_{VD}, \rho_{FD}, \rho_{CD}$		−0.111***	−0.037	0.078*
		(0.032)	(0.029)	(0.038)
ρ_{VF}, ρ_{VC}			0.453***	−0.418***
			(0.048)	(0.049)
ρ_{FC}				0.128*
				(0.058)
Total effect routine patients				
Vol	0.052			
	(0.061)			
Foc	−0.201***			
	(0.060)			
Con	−0.070			
	(0.073)			
Total effect complex patients				
Vol	0.164**			
	(0.063)			
Foc	0.027			
	(0.057)			
Con	−0.228***			
	(0.072)			
Effect differences				
Δ Vol (PR, PC)	−0.111*			
	(0.051)			
Δ Foc (PR, PC)	−0.228***			
	(0.043)			
Δ Con (PR, PC)	0.158***			
	(0.047)			
Observations	265,133			
Segments-in-hospitals	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 8 Simultaneous equations models for seven-day mortality: K=3

Mortality equation	Simultaneous equations model		
Vol	0.034		
	(0.055)		
Vol * PR	0.016		
	(0.041)		
Vol * PC	0.127**		
	(0.041)		
Foc	−0.003		
	(0.053)		
Foc * PR	−0.189***		
	(0.037)		
Foc * PC	0.038		
	(0.033)		
Con	−0.136+		
	(0.072)		
Con * PR	0.089*		
	(0.039)		
Con * PC	−0.068*		
	(0.034)		
Selection equations (Ivs)		Vol	Foc	Con
DD_V, DD_F, DD_C		−0.023***	−0.002	−0.003
		(0.005)	(0.006)	(0.005)
$D1_V, D1_F, D1_C$		0.461***	0.606***	0.316***
		(0.080)	(0.098)	(0.085)
$D2_V, D2_F, D2_C$		0.130*	0.229**	0.141*
		(0.063)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$		0.138*	0.116	0.164**
		(0.057)	(0.077)	(0.062)
Error correlations				
$\rho_{VD}, \rho_{FD}, \rho_{CD}$		−0.108***	−0.043	0.063
		(0.031)	(0.030)	(0.041)
ρ_{VF}, ρ_{VC}			0.462***	−0.405***
			(0.049)	(0.051)
ρ_{FC}				0.138*
				(0.059)
Total effect routine patients				
Vol	0.049			
	(0.061)			
Foc	−0.192**			
	(0.061)			
Con	−0.047			
	(0.077)			
Total effect complex patients				
Vol	0.161*			
	(0.063)			
Foc	0.035			
	(0.057)			
Con	−0.205**			
	(0.079)			
Effect differences				
Δ Vol (PR, PC)	−0.111*			
	(0.051)			
Δ Foc (PR, PC)	−0.227***			
	(0.043)			
Δ Con (PR, PC)	0.158***			
	(0.048)			
Observations	265,133			
Segments-in-hospitals	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 9 Simultaneous equations models for seven-day mortality: K=4

Mortality equation	Simultaneous equations model
Vol	0.035
	(0.056)
Vol * PR	0.016
	(0.041)
Vol * PC	0.127**
	(0.041)
Foc	−0.006
	(0.053)
Foc * PR	−0.189***
	(0.037)
Foc * PC	0.038
	(0.033)
Con	−0.142*
	(0.071)
Con * PR	0.089*
	(0.039)
Con * PC	−0.068*
	(0.034)

Selection equations (Ivs)	Vol	Foc	Con
DD_V, DD_F, DD_C	−0.023***	−0.003	−0.002
	(0.005)	(0.006)	(0.005)
$D1_V, D1_F, D1_C$	0.460***	0.605***	0.324***
	(0.080)	(0.097)	(0.085)
$D2_V, D2_F, D2_C$	0.128*	0.225**	0.144*
	(0.063)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$	0.134*	0.114	0.164**
	(0.057)	(0.077)	(0.062)
$D4_V, D4_F, D4_C$	−0.060	−0.071	0.148*
	(0.053)	(0.065)	(0.068)
Error correlations			
$\rho_{VD}, \rho_{FD}, \rho_{CD}$	−0.109***	−0.041	0.067+
	(0.031)	(0.030)	(0.040)
ρ_{VF}, ρ_{VC}		0.462***	−0.407***
		(0.049)	(0.051)
ρ_{FC}			0.136*
			(0.059)

Total effect routine patients	
Vol	0.051
	(0.061)
Foc	−0.196**
	(0.062)
Con	−0.053
	(0.076)
Total effect complex patients	
Vol	0.162*
	(0.063)
Foc	0.032
	(0.058)
Con	−0.210**
	(0.078)
Effect differences	
Δ Vol (PR, PC)	−0.111*
	(0.051)
Δ Foc (PR, PC)	−0.228***
	(0.043)
Δ Con (PR, PC)	0.158***
	(0.048)
Observations	265,133
Segments-in-hospitals	2,067

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 10 Simultaneous equations models for seven-day mortality: Parsimonious model with K=1 and no DD

Mortality equation	Simultaneous equations model		
Vol	0.041		
	(0.057)		
Vol * PR	0.015		
	(0.041)		
Vol * PC	0.127**		
	(0.041)		
Foc	−0.006		
	(0.054)		
Foc * PR	−0.189***		
	(0.037)		
Foc * PC	0.038		
	(0.033)		
Con	−0.156*		
	(0.076)		
Con * PR	0.089*		
	(0.039)		
Con * PC	−0.069*		
	(0.034)		
Selection equations (Ivs)		Vol	Foc	Con
$D1_V, D1_F, D1_C$		0.610***	0.610***	0.331***
		(0.064)	(0.064)	(0.067)
Error correlations				
$\rho_{VD}, \rho_{FD}, \rho_{CD}$		−0.113***	−0.041	0.076+
		(0.032)	(0.031)	(0.044)
ρ_{VF}, ρ_{VC}			0.461***	−0.404***
			(0.048)	(0.051)
ρ_{FC}				0.143*
				(0.060)
Total effect routine patients				
Vol	0.056			
	(0.063)			
Foc	−0.194**			
	(0.063)			
Con	−0.067			
	(0.083)			
Total effect complex patients				
Vol	0.168**			
	(0.065)			
Foc	0.033			
	(0.060)			
Con	−0.225**			
	(0.083)			
Effect differences				
Δ Vol (PR, PC)	−0.112*			
	(0.051)			
Δ Foc (PR, PC)	−0.227***			
	(0.043)			
Δ Con (PR, PC)	0.158***			
	(0.048)			
Observations	265,133			
Segments-in-hospitals	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

6. Evidence for hospital selection using observed variables

Hospital studies of the mortality effects of volume and focus are rightly concerned with the possibility that a positive effect of a hospital's high volume or high focus on a disease segment may be a consequence of the hospital's ability to cherry-pick healthier patients rather than its operational superiority. It is therefore necessary to control for unobserved factors that affect hospital selection, as we do in the simultaneous equations model in the main paper. In this section, we provide some descriptive evidence that this is necessary by showing that patients' mortality risks differ between by independent variable on observed factors. It is therefore likely that there are additional unobserved selection factors. To this end, we estimate a probit model for mortality at the patient level, using patient characteristics only (disease segment, age, gender, emergency admission status, Elixhauser comorbidities) and then test whether the average model-predicted mortality risk differs between high-volume (focus / concentration) and low-volume (focus / concentration) hospitals by regressing the model-predicted mortality risk onto the respective binary variables. Table 11 summarizes the results and shows clear differences in mortality risks for the three variables of interest and suggest the presence of endogeneity.

Table 11 95% Confidence intervals for patient-risk adjusted seven-day mortality rate by hospital type

Variable	High	Low
Volume	[2.78%, 2.81%]	[4.10%, 4.19%]
Focus	[2.72%, 2.76%]	[3.59%, 3.66%]
Concentration	[3.32%, 3.38%]	[2.80%, 2.84%]

7. Routing process

7.1. Patient flow through departments

In this section we explain the routing processes to hospital departments in some more detail. Since transfers between general surgery and internal medicine approximately account for 50% of the transfers, we limit our illustration to these two transfer types. The patient flow in our data, including seven-day transfer rates, are indicated in Figures 3 and 4. Emergency and elective patients differ in their initially assigned department. While the percentage of patients admitted to general surgery is comparable (20.8% vs. 19.4%), emergency patients are more often admitted to internal medicine (44.0%) than other department types (35.2%), while elective patients are more often admitted to other departments (53.7%), e.g. cardiology and cardio-surgery, than internal medicine (26.9%). Note, however, that the indicated likelihoods of transfer within the first seven days are of similar magnitudes between electives and emergencies, with a small tendency for more transfers of emergency patients.

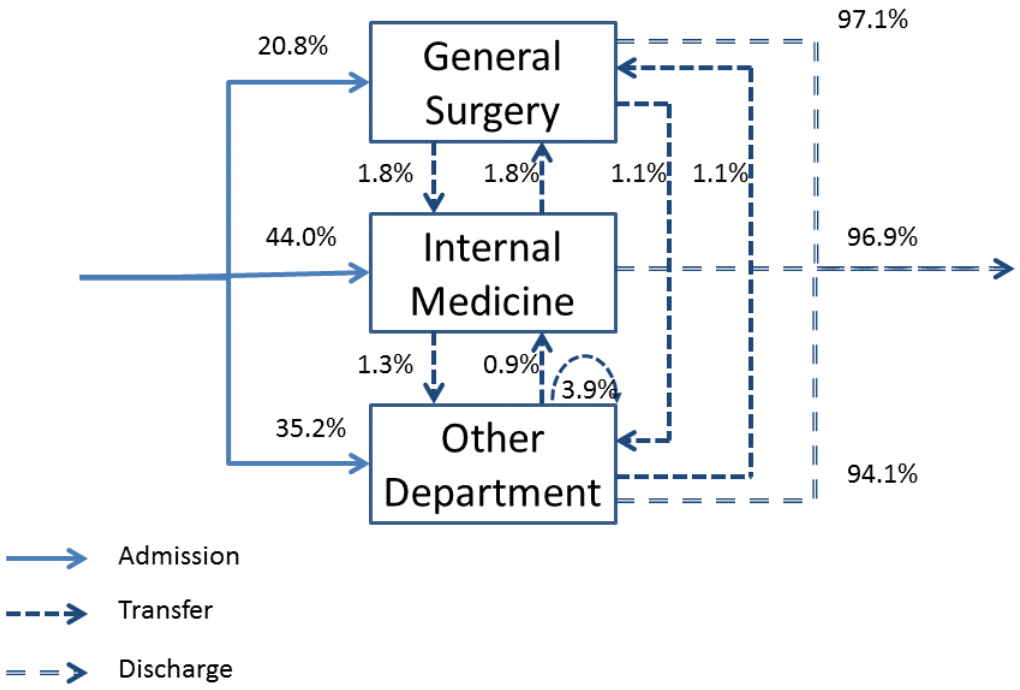


Figure 3 **Emergency patient flow**

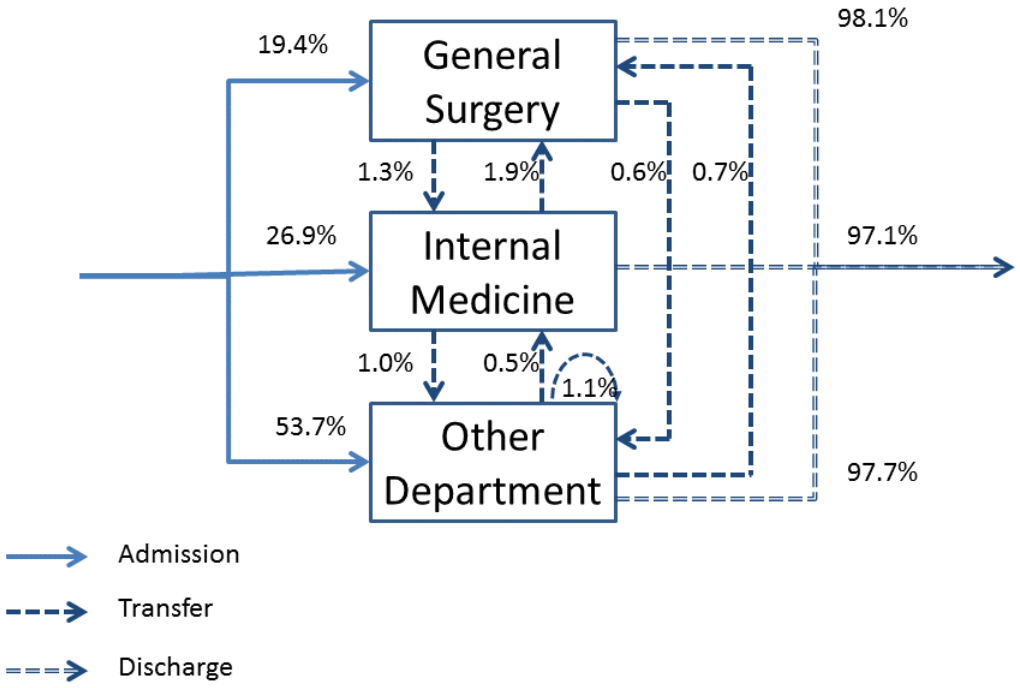


Figure 4 **Elective patient flow**

We would have liked to be able to explore the reasons for transfers of elective patients in more detail but such information is not available in our data. Nevertheless, some observed differences between the transferred and non-transferred may be indicative of some of the transfer reasons. Table 12 outlines descriptive statistics for elective patients who were initially admitted to general surgery and subsequently transferred to internal medicine (first column) or not transferred to internal medicine (either staying in general surgery or transferred to another department). The transfer group is on average older and more often complex than their non-transferred counterparts. The most striking difference relates to surgical procedures: Patients transferred from general surgery to internal medicine receive a surgical procedure in 29% of the cases compared to 72% of the patients not transferred to internal medicine. This suggests that patients are often transferred from general surgery to internal medicine because they are not sufficiently healthy to undergo a planned surgery (which is more likely to be the case for complex and elderly patients). These allocations to the general surgery department may be deemed as departmental allocation errors.

Table 13 outlines descriptive statistics for elective patients who were initially assigned to internal medicine and subsequently transferred to general surgery (first column) or not transferred to general surgery. While there is less of a difference in patient demographics, the notable difference again exists with respect to surgical procedures. Patients who get transferred to general surgery receive in 82% of the cases a surgical procedure compared to 6% of the patients not transferred to general surgery. This suggests that patients are often transferred from internal medicine to general surgery when a reevaluation of their condition suggests that surgery is preferable to conventional treatment. Again, this can be regarded as an departmental allocation error.

Table 12 **Elective patient flow drivers: Transfers from general surgery to internal medicine**

Variable	Transfer to IM N=623	No Transfer to IM N=49,267	Difference
Age	69.3	59.2	10.1
Complex	50%	19%	31%
Average day of Transfer	2.6	NA	
Surgery	29%	72%	-43%

Table 13 **Elective patient flow drivers: Transfers from internal medicine to general surgery**

Variable	Transfer to GS N=1,327	No Transfer to GS N=67,637	Difference
Age	67.2	68.4	-1.20
Complex	42%	36%	6%
Average day of Transfer	3.0	NA	
Surgery	82%	6%	76%

7.2. Segment concentration and allocation errors: A mathematical model

In this section, we present a simple mathematical model that illustrates why patients whose hospitals route a higher proportion of their patient segment in a single "default department" have fewer departmental allocation errors and that this effect is stronger for more complex patients. The paper complements the arguments put forward in Section 3.3 of the main paper.

Suppose a gatekeeper has to allocate newly arriving customers to one of two departments, A or B and that A is the most appropriate department for most but not all customers. When a new customer arrives, the gatekeeper will always perform a test to confirm that A is indeed the best department for the customer. If the test result is positive (event τ_A^1) then the gatekeeper will refer the customer to department A. If, however, the test is negative, suggesting that department B is more appropriate (event τ_B^1), the gatekeeper has two conflicting pieces of information: a prior belief that A is the most appropriate department and a conflicting test results (event τ_B^1). To keep the mathematics simple, we assume that the gatekeeper resolves the tie by performing a second, decisive test: If the result of this second test contradicts the first test and suggests that A is the appropriate department (event τ_A^2), the gatekeeper will ignore the contradictory first test result and send the customer to A. If, however, the second test confirms the first test result, suggesting that B is more appropriate (event τ_B^2), then the gatekeeper overrules the larger prior for A and sends the customer to B.

This simple gatekeeping process reflects practice fairly well and allows for a straightforward mathematical justification that, under reasonable assumptions about the relative specificity and sensitivity of the two test, an increase in segment concentration in default department A reduces department allocation errors and that this effect is more pronounced for more complex patients. We will indicate by A and B the events that department A or B, respectively, is the most appropriate department for a sampled customer, and by $P_A, P_B = (1 - P_A)$ the corresponding probabilities. Part 1 of the proposition provides conditions that ensure that the total error rate of the gatekeeping process decreases with increasing P_A, while Part 2 provides conditions that ensure that this statement extends to the proportion P_R of customers that get routed into department A. Finally, Part 3, provides conditions that ensure that the rate of decrease is larger for more complex patients.

PROPOSITION 1. *Let* $\alpha = \min\{P(\tau_A^1 \mid A), P(\tau_A^2 \mid A, \tau_B^1)\}$ *and* $\beta = \max\{P(\tau_B^1 \mid B), P(\tau_B^2 \mid B, \tau_B^1)\}$.

1. *If* $\alpha \geq 1 - \sqrt{1 - \beta^2}$, *then the error probability of the described gatekeeping process decreases monotonically with in* P_A.

2. *If* $\sqrt{3} - 1 \leq \alpha_i, \beta_i \leq 1$ *then the proportion* P_R *of customers that get routed into department* A *increases monotonically with* P_A.

3. *Let* $\alpha_1(c) = P(\tau_A^1 \mid A), \alpha_2(c) = P(\tau_A^2 \mid A, \tau_B^1), \beta_1(c) = P(\tau_B^1 \mid B), \beta_2(c) = P(\tau_B^2 \mid B, \tau_B^1)$ *be monotonically decreasing functions of a measure* c *of patient complexity. If* $\alpha_i(c) \geq 0.5$, $\beta_i(c) \geq 0.5$ *and* $\beta_i'(c) \leq \alpha_i'(c) \leq 0$ *for all* c*, then the error probability of the described gatekeeping process decreases more strongly in the probability* P_A *for more complex patients, provided the condition of Part 1 holds. Since* P_R *is an increasing linear function of* P_A*, the same statement holds for* P_R*, provided the condition of Part 2 holds.*

Proof. Let $\alpha_1 = P(\tau_A^1 \mid A), \alpha_2 = P(\tau_A^2 \mid A, \tau_B^1), \beta_1 = P(\tau_B^1 \mid B), \beta_2 = P(\tau_B^2 \mid B, \tau_B^1)$. The routing errors of the described gatekeeping process correspond to three event combinations $(\tau_A^1, B), (\tau_B^1, \tau_A^2, B), (\tau_B^1, \tau_B^2, A)$ with associated probabilities

$$\begin{aligned} P(\tau_A^1, B) &= P_B P(\tau_A^1 \mid B) = (1 - P_A)(1 - \beta_1) \\ P(\tau_B^1, \tau_A^2, B) &= P_B(\tau_B^1 \mid B) P(\tau_A^2 \mid B, \tau_B^1) = (1 - P_A)\beta_1(1 - \beta_2) \\ P(\tau_B^1, \tau_B^2, A) &= P_A P(\tau_B^1 \mid A) P(\tau_B^2 \mid A, \tau_B^1) = P_A(1 - \alpha_1)(1 - \alpha_2). \end{aligned}$$

Summing up the three terms gives the total error probability

$$P_E = (1 - \beta_1\beta_2)(1 - P_A) + (1 - \alpha_1)(1 - \alpha_2)P_A,$$

which is a monotonically decreasing function of P_A as long as $(1 - \beta_1\beta_2) \geq (1 - \alpha_1)(1 - \alpha_2)$. The latter inequality follows from the proposition's assumption $\alpha \geq 1 - \sqrt{1 - \beta^2}$ (which implies $(1 - \beta^2) \geq (1 - \alpha)^2$) and the fact that $\alpha \leq \alpha_i$ and $\beta \geq \beta_i$ for $i = 1, 2$.

To see that Part 2 holds, we calculate the proportion P_R of customers routed into department A by the gatekeeping process

$$\begin{aligned} P_R &= P(\tau_A^1, A) + P(\tau_A^1, B) + P(\tau_B^1, \tau_A^2, A) + P(\tau_B^1, \tau_A^2, B) \\ &= P_A P(\tau_A^1 \mid A) + P_B P(\tau_A^1 \mid B) + P_A P(\tau_B^1 \mid A) P(\tau_A^2 \mid A, \tau_B^1) + P_B P(\tau_B^1 \mid B) P(\tau_A^2 \mid B, \tau_B^1) \\ &= P_A\alpha_1 + (1 - P_A)(1 - \beta_1) + P_A(1 - \alpha_1)\alpha_2 + (1 - P_A)\beta_1(1 - \beta_2). \end{aligned} \tag{7}$$

The coefficient of P_A is $\alpha_1 + \alpha_2 - \alpha_1\alpha_2 + \beta_1\beta_2 - 1$ which, under the assumption $\sqrt{3} - 1 \leq \alpha_i, \beta_i \leq 1$, is bounded below by $(\sqrt{3} - 1) + (\sqrt{3} - 1) - 1 + (3 - 2\sqrt{3} + 1) - 1 = 0$

To see Part 3, recall from Part 1 and Part 2 that the total error probability P_E is of the form $P_E = g(c) + f(c)P_A$, and $P_R = h(c) + f_R(c)P_A$. Combining these relations, this yields

$$P_E = g(c) + f(c)\frac{P_R - h(c)}{f_R(c)} = g(c) - \frac{f(c)h(c)}{f_R(c)} + P_R\frac{f(c)}{f_R(c)}.$$

Then our claim holds if $(\frac{f(c)}{f_R(c)})' = \frac{f'(c)f_R(c) - f(c)f_R'(c)}{f_R(c)^2} \leq 0$. Since $f(c) \leq 0$ and $f_R(c) \geq 0$, the inequality holds if $f'(c) \leq 0$ and $f_R'(c) \leq 0$. Note that

$$f'(c) = -\alpha_1'(c) - \alpha_2'(c) + \alpha_1'(c)\alpha_2(c) + \alpha_1(c)\alpha_2'(c) + \beta_1'(c)\beta_2(c) + \beta_1(c)\beta_2'(c)$$

$$\begin{aligned}
&\leq -\alpha_1'(c) - \alpha_2'(c) + \alpha_1'(c)\alpha_2(c) + \alpha_1(c)\alpha_2'(c) + \alpha_1'(c)\beta_2(c) + \beta_1(c)\alpha_2'(c) \\
&= \alpha_1'(c)(\alpha_2(c) + \beta_2(c) - 1) + \alpha_2'(c)(\alpha_1(c) + \beta_1(c) - 1) \\
&\leq 0.
\end{aligned}$$

To see that $f_R'(c) \leq 0$ note that $f_R(c)$ is the coefficient of P_R in equation (7) and hence of the form $f_R(c) = \alpha_1(c) + \alpha_2(c) - \alpha_1(c)\alpha_2(c) + \beta_1(c)\beta_2(c) - 1$. Therefore

$$\begin{aligned}
f_R'(c) &= \alpha_1'(c) + \alpha_2'(c) - \alpha_1'(c)\alpha_2(c) - \alpha_1(c)\alpha_2'(c) + \beta_1'(c)\beta_2(c) + \beta_1(c)\beta_2'(c) \\
&= \alpha_1'(c)(1 - \alpha_2(c)) + \alpha_2'(c)(1 - \alpha_1(c)) + \beta_1'(c)\beta_2(c) + \beta_1(c)\beta_2'(c).
\end{aligned}$$

The latter term is non-positive because, by assumption, $\alpha_1'(c), \alpha_2'(c), \beta_1'(c), \beta_2'(c) \leq 0$ and $0 \leq \alpha_1(c), \alpha_2(c), \beta_1(c), \beta_2(c) \leq 1$. □

Note that the condition for Part 1, $\alpha \geq 1 - \sqrt{1 - \beta^2}$, is significantly weaker than $\alpha \geq \beta$. In fact, the function $f(\beta) = 1 - \sqrt{1 - \beta^2}$ is convex and increases from 0 to 1 as β ranges from 0 to 1, with $f'(0) = 0$ and $f'(1) = \infty$. Part 2 makes an assumption that the true positive rate exceeds 0.73. As for Part 3, the assumptions $\alpha_i, \beta_i \geq 0.5$ (implied by the assumption of Part 2) and $\alpha_i'(c), \beta_i'(c) \leq 0$ are uncontroversial: One would not want to use a test with a false positive rate above 50% and it seems sensible to assume that the identification of the correct department in the various testing stages is more difficult for more complex patients. The assumption $\beta_i'(c) \leq \alpha_i'(c)$ requires that for those patients who should not be routed to the segment's default department A, the difficulty of identifying the correct department with a test declines more rapidly with increasing patient complexity than for those patients who should be routed to the default department.

7.3. Effect of segment concentration on departmental transfers

In our theory section, we claim that a higher degree of segment concentration reduces departmental routing errors, in particular for complex patients, who require more time for an accurate departmental allocation. This section provides empirical support for this claim. We consider internal transfers, specifically transfers from the admitting department to another hospital department within seven days of hospital admission, as indicative of department allocation errors; later transfers are more likely to be part of standard care pathways, as in the case of transfer to a rehabilitation department. The dependent variable is coded as a binary variable with 1 indicating a departmental transfer within seven days of the patient's hospital admission. We discard transfers to intensive care units as they are not indicative of an allocation error and instead reflect a deterioration in a patient's health. Since not all hospitals provided information about transfers in their standardized discharge records, we restricted the transfer sample to the 56 hospitals that provided this information. We

further restricted the sample to segments with sufficient variance in transfer rates by excluding all segments with a seven-day transfer rate below 1%. In addition, we exclude perfect predictors and these exclusions left us with a seven-day transfer subsample of 383,628 patient episodes in 52 hospitals and 98 disease segments.

We use all control variables defined in the Section Control Variables in the main paper. While concerns that the relationship between segment concentration and transfers is confounded by factors related to the hospital structure, specifically the number of departments and therefore the number of transfer alternatives, are mitigated by incorporating hospital fixed effects, transfer policies may well differ between segments within the same hospital. Therefore, we control for the general transfer rate in segment s's default department d in hospital h by calculating the transfer rate for all patients $i \in d$ outside of segment s. We estimate a probit model and additionally test by means of a biprobit model whether concentration is endogeneous, i.e. whether patients admitted to a high-concentration hospital differ in their unobservable transfer risk from patients admitted to a low-concentration hospital.

The probit and biprobit estimation results for the concentration–transfer relationship are provided in Table 14. We find no evidence for endogeneity for the concentration-transfer relationship in a recursive biprobit model ($\rho = -0.160, p > 0.1$) and we can rely on the more efficient probit estimates, which are shown in Table 14. These results show that hospitals with a higher segment concentration have a lower transfer rate for benchmark patients ($\tilde{\beta} = -0.253, p < 0.001$) and that this beneficial concentration effect is further amplified for complex patients ($\tilde{\beta} = -0.157, p < 0.001$) and reduced for routine patients ($\tilde{\beta} = +0.100, p < 0.001$). While it is weaker, the overall concentration effect remains significant for routine patients ($-0.253 + 0.100 = -0.153, p < 0.001$). Table 15 shows that the effect sizes are operationally significant, and that the difference in estimated transfer rates between hospitals with a low and high segment concentration is strongest for complex patients.

8. Sample selection thresholds and in-hospital observation period

8.1. Varying exclusion criteria

The sample used for the main paper was trimmed to increase its homogeneity, as shown in Figure 5. The exclusion criteria include several discretionary thresholds (number of hospital patients, number of departmental patients) and we therefore checked the robustness of the results by varying the organizational thresholds (hospital, department) between $(0, 0)$ (no exclusion), $(1000, 50)$ (sample in main paper), and $(2500, 100)$ (increased trimming). At the same time we varied the mortality thresholds for inclusion of segments between 0% (no exclusion), 1% (sample in main paper), and 2%. We report the analyses for the largest (Table 16) and the smallest sample (Table 17) and the

Table 14 Probit and biprobit coefficient estimates for seven-day departmental transfer

Variables	(1) Transfer	(2) Transfer
Con	−0.253***	0.010
	(0.028)	(0.272)
Con * PR	0.100***	0.100***
	(0.030)	(0.030)
Con * PC	−0.157***	−0.157***
	(0.040)	(0.040)
Total effect routine patients		
Con	−0.153***	0.110
	(0.034)	(0.272)
Total effect complex patients		
Con	−0.410***	−0.147
	(0.043)	(0.276)
Effect differences		
Δ Con (PR, PC)	0.257***	0.257***
	(0.049)	(0.049)
Selection equation		Con
DD_C		−0.012**
		(0.004)
$D1_C$		0.231***
		(0.056)
$D2_C$		0.172***
		(0.043)
$D3_C$		0.048
		(0.040)
$D4_C$		0.012
		(0.041)
$D5_C$		0.022
		(0.038)
ρ_{CT}		−0.160
		(0.158)
Observations	383,628	383,628
Cluster	4,395	4,395

Clustered standard errors in parentheses. The model includes the control variables as per Table 2 and the transfer rate of patients outside the focal segment. *** p<0.001, ** p<0.01, * p<0.05

Table 15 Estimated risk-adjusted seven-day in-hospital transfer rates for concentration

	Concentration–transfer		
	Low	High	p-value
Routine	2.00%	1.41%	<0.001
Benchmark	4.01%	2.39%	<0.001
Complex	7.48%	3.56%	<0.001

results of the other combinations are quite similar: For each of the eight combinations, we obtain consistent results for our focus-hypotheses and concentration-hypothesis that are at least weakly significant. The volume-hypothesis finds partial support unless the mortality threshold is increased to 2%, when statistical power appears to too low to identify significant effects. Note that the signs of the estimated volume effects (Total effect routine patients / Total effect complex patients / Effect difference) in the simultaneous equations model are the same as in Table 1 of the main

paper.

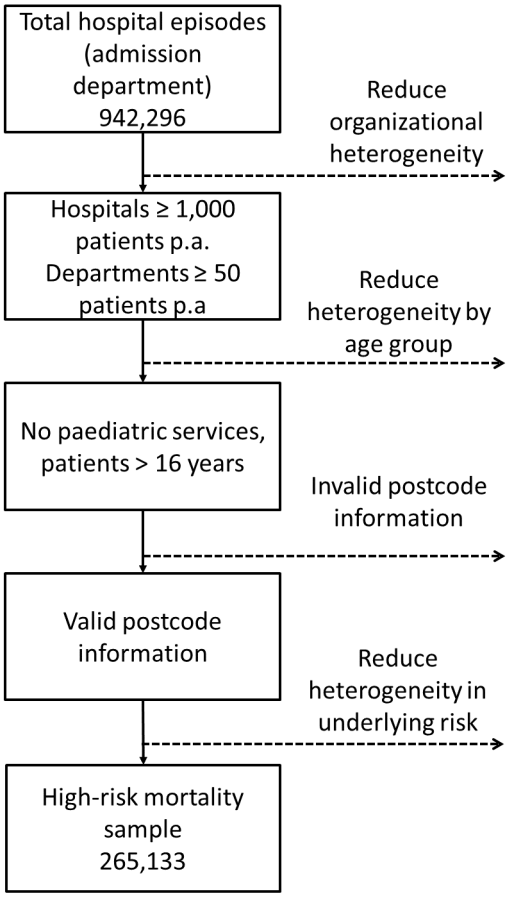


Figure 5 Construction of data sample

8.2. Expansion of observation period

We expanded the observation period by not only focusing on mortality for the first seven days but also considered events up to day 8, 9, 10 and finally all inpatient deaths irrespective of their timing. The results concerning the total effects and differential effect between routine and complex patients remain comparable to the seven-day window reported in the main paper. Table 18 shows the results of all inpatient deaths.

9. Patient complexity

9.1. Varying comorbidity thresholds for complex patients

In the main paper, we had reported results where complex patients were defined as emergency patients with at least three Elixhauser comorbidities. We replicate the analysis using two comorbidities and four comorbidities as the threshold, with results in Tables 19 and 20. The total effects

Table 16 Probit and simultaneous equations models for seven-day mortality: largest sample

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.065∗	0.035		
	(0.029)	(0.052)		
Vol * PR	−0.026	−0.022		
	(0.037)	(0.036)		
Vol * PC	0.088∗	0.086∗		
	(0.038)	(0.038)		
Foc	−0.072∗∗∗	0.059		
	(0.021)	(0.047)		
Foc * PR	−0.178∗∗∗	−0.168∗∗∗		
	(0.033)	(0.033)		
Foc * PC	0.048	0.045		
	(0.031)	(0.031)		
Con	−0.053∗	−0.089		
	(0.022)	(0.067)		
Con * PR	0.087∗∗	0.086∗∗		
	(0.033)	(0.033)		
Con * PC	−0.053+	−0.052		
	(0.032)	(0.032)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.026∗∗∗	−0.014∗∗∗	−0.011∗∗
			(0.003)	(0.003)	(0.003)
$D1_V, D1_F, D1_C$			0.345∗∗∗	0.418∗∗∗	0.205∗∗∗
			(0.043)	(0.058)	(0.051)
$D2_V, D2_F, D2_C$			0.153∗∗∗	0.172∗∗∗	0.184∗∗∗
			(0.036)	(0.047)	(0.038)
$D3_V, D3_F, D3_C$			0.121∗∗∗	0.098∗	0.075∗
			(0.033)	(0.045)	(0.037)
$D4_V, D4_F, D4_C$			−0.063∗	0.015	0.015
			(0.031)	(0.040)	(0.038)
$D5_V, D5_F, D5_C$			0.006	−0.014	0.012
			(0.031)	(0.043)	(0.036)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.086∗∗	−0.089∗∗	0.021
			(0.027)	(0.027)	(0.039)
ρ_{VF}, ρ_{VC}				0.454∗∗∗	−0.230∗∗∗
				(0.027)	(0.032)
rho_{FC}					0.184∗∗∗
					(0.035)
Total effect routine patients					
Vol	−0.091∗	0.014			
	(0.036)	(0.056)			
Foc	−0.250∗∗∗	−0.109∗			
	(0.028)	(0.053)			
Con	0.034	−0.002			
	(0.030)	(0.068)			
Total effect complex patients					
Vol	0.023	0.122∗			
	(0.039)	(0.058)			
Foc	−0.024	0.104∗			
	(0.029)	(0.050)			
Con	−0.106∗∗∗	−0.140+			
	(0.031)	(0.072)			
Effect differences					
Δ Vol (PR, PC)	−0.114∗	−0.108∗			
	(0.046)	(0.045)			
Δ Foc (PR, PC)	−0.226∗∗∗	−0.213∗∗∗			
	(0.038)	(0.038)			
Δ Con (PR, PC)	0.140∗∗∗	0.138∗∗∗			
	(0.041)	(0.041)			
Observations	653,367	653,367			
Segments-in-hospitals	6,095	6,095			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 17 Probit and simultaneous equations models for seven-day mortality: smallest sample

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.143∗∗∗	−0.008		
	(0.039)	(0.062)		
Vol * PR	0.097∗	0.100∗		
	(0.047)	(0.046)		
Vol * PC	0.130∗∗	0.128∗∗		
	(0.047)	(0.047)		
Foc	−0.050+	0.051		
	(0.028)	(0.061)		
Foc * PR	−0.214∗∗∗	−0.206∗∗∗		
	(0.046)	(0.045)		
Foc * PC	0.031	0.031		
	(0.040)	(0.040)		
Con	−0.070∗	−0.149+		
	(0.030)	(0.086)		
Con * PR	0.127∗∗	0.123∗∗		
	(0.047)	(0.047)		
Con * PC	−0.022	−0.019		
	(0.042)	(0.041)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.026∗∗∗	0.004	0.007
			(0.006)	(0.007)	(0.008)
$D1_V, D1_F, D1_C$			0.494∗∗∗	0.509∗∗∗	0.370∗∗∗
			(0.110)	(0.108)	(0.104)
$D2_V, D2_F, D2_C$			0.293∗∗∗	0.296∗∗	0.175∗
			(0.081)	(0.094)	(0.085)
$D3_V, D3_F, D3_C$			0.206∗∗	0.072	0.199∗
			(0.074)	(0.088)	(0.083)
$D4_V, D4_F, D4_C$			−0.060	0.008	0.134
			(0.070)	(0.078)	(0.089)
$D5_V, D5_F, D5_C$			0.068	0.023	0.008
			(0.067)	(0.087)	(0.084)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.113∗∗∗	−0.073∗	0.058
			(0.034)	(0.036)	(0.051)
ρ_{VF}, ρ_{VC}				0.471∗∗∗	−0.332∗∗∗
				(0.067)	(0.071)
ρ_{FC}					0.157∗
					(0.078)
Total effect routine patients					
Vol	−0.046	0.092			
	(0.050)	(0.069)			
Foc	−0.264∗∗∗	−0.158∗			
	(0.041)	(0.064)			
Con	0.057	−0.027			
	(0.042)	(0.095)			
Total effect complex patients					
Vol	−0.013	0.119+			
	(0.049)	(0.072)			
Foc	−0.019	0.081			
	(0.038)	(0.065)			
Con	−0.092∗	−0.168+			
	(0.041)	(0.093)			
Effect differences					
Δ Vol (PR, PC)	−0.034	−0.027			
	(0.058)	(0.058)			
Δ Foc (PR, PC)	−0.245∗∗∗	−0.237∗∗∗			
	(0.051)	(0.051)			
Δ Con (PR, PC)	0.149∗∗∗	0.142∗∗			
	(0.055)	(0.055)			
Observations	139,028	139,028			
Segments-in-hospitals	904	904			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 18 Probit and simultaneous equations models for complete inpatient mortality

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.055+	0.083		
	(0.030)	(0.051)		
Vol * PR	−0.057	−0.052		
	(0.036)	(0.035)		
Vol * PC	0.089**	0.087*		
	(0.034)	(0.034)		
Foc	−0.073***	−0.009		
	(0.021)	(0.050)		
Foc * PR	−0.170***	−0.164***		
	(0.032)	(0.032)		
Foc * PC	0.071*	0.069*		
	(0.030)	(0.029)		
Con	−0.055*	−0.115+		
	(0.021)	(0.060)		
Con * PR	0.065+	0.061+		
	(0.034)	(0.034)		
Con * PC	−0.046	−0.043		
	(0.029)	(0.029)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023***	−0.004	−0.004
			(0.004)	(0.005)	(0.005)
$D1_V, D1_F, D1_C$			0.442***	0.540***	0.330***
			(0.066)	(0.084)	(0.073)
$D2_V, D2_F, D2_C$			0.121*	0.217**	0.158**
			(0.052)	(0.066)	(0.055)
$D3_V, D3_F, D3_C$			0.124*	0.145*	0.128*
			(0.049)	(0.067)	(0.052)
$D4_V, D4_F, D4_C$			−0.062	−0.016	0.106+
			(0.045)	(0.057)	(0.057)
$D5_V, D5_F, D5_C$			0.007	−0.014	−0.033
			(0.045)	(0.065)	(0.051)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.105***	−0.049+	0.050
			(0.027)	(0.028)	(0.034)
ρ_{VF}, ρ_{VC}				0.452***	−0.412***
				(0.042)	(0.043)
ρ_{FC}					0.136**
					(0.051)
Total effect routine patients					
Vol	−0.111**	0.031			
	(0.037)	(0.055)			
Foc	−0.243***	−0.173**			
	(0.031)	(0.058)			
Con	0.010	−0.054			
	(0.031)	(0.063)			
Total effect complex patients					
Vol	0.034	0.170**			
	(0.037)	(0.056)			
Foc	−0.002	0.060			
	(0.030)	(0.054)			
Con	−0.101***	−0.157*			
	(0.029)	(0.063)			
Effect differences					
Δ Vol (PR, PC)	−0.146***	−0.139***			
	(0.044)	(0.044)			
Δ Foc (PR, PC)	−0.241***	−0.233***			
	(0.041)	(0.040)			
Δ Con (PR, PC)	0.111*	0.104*			
	(0.043)	(0.043)			
Observations	329,424	329,424			
Segments-in-hospitals	2,402	2,402			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10

for routine and complex patients as well as the effect difference between routine and complex patients remain comparable to those reported in the main paper, i.e. we find partial support for our volume-outcome hypothesis and our hypothesis concerning focus and concentration are fully supported.

9.2. An alternative patient complexity measure

Our measurement of patient complexity in the main paper is based on the number of Elixhauser comorbidities registered in the patient's discharge record. This measure is coarse and does not take into account that the complicating effect of specific comorbidities differ from patient segment to patient segment. Therefore simple addition leads to measurement errors. Also, there are non-Elixhauser comorbidities that could lead to increased patient complexity. As a robustness check we report in this section the results for a different patient complexity classification, akin to the complexity measure used in Clark (2012). The measure is based on information extracted from a patient's diagnosis related group (DRG), which classify patients by conditions or procedures and are used for reimbursement of hospital services and therefore available in standardized discharge records. First, like Clark (2012), the DRG complexity measure is based on all actual secondary diagnoses in the discharge record. In the German system, every secondary diagnosis obtains a CC score (CCL Wert), ranging from 0 (no comorditidy or complexity) to 4 (extremely severe comorbidity and complexity). The CC score of a secondary diagnosis is set to 0 if there is a close connection with the main diagnosis (accordingly to Clark's requirement that a comorbidity should be a secondary diagnosis that falls into a different disease category from the primary diagnosis). The German DRG system then calculates an aggregate CC score for the patient, a patient-level complexity score (PCCL Wert) between 0 (no relevant comorbidity and complexity) and 4 (most severe comorbidity and complexity). If different PCCL levels are not associated with different cost implications, then the DRG codes are not subdivided by PCCL level (indicated by Z as the 4th digit), otherwise the four levels are included as letters A-D in 4th digit of the DRG code. We believe this is very close to Clark (2012) and the best we can do to get close to Clarks manual method with our data. We should also stress that this is a nationally agreed complexity score in the German hospital system.

We use this classification to identify complex and routine patients within each DRG. Specifically, we replace the comorbidity criterion "more than 3 comorbidities" by the criterion "DRG classification A", which refers to the fourth letter of the code, where "A" always refers to the highest complication level. We classify patients as complex if they are emergency admissions and the DRG CC classification letter is A. Note that DRGs are only subdivided if differences in CC levels have significant cost implications. When DRGs are not subdivided, we assign them to benchmark patients if they are emergency admissions and to routine patients if they are elective admissions.

Table 19 Probit and simultaneous equations models for seven-day mortality: 2 comorbidity threshold

Mortality equation	Probit	Simultaneous equations model
Vol	−0.089**	0.059
	(0.034)	(0.056)
Vol * PR	−0.079	−0.072
	(0.053)	(0.052)
Vol * PC	0.068*	0.065+
	(0.034)	(0.034)
Foc	−0.130***	−0.077
	(0.023)	(0.054)
Foc * PR	−0.089+	−0.084+
	(0.048)	(0.048)
Foc * PC	0.126***	0.124***
	(0.030)	(0.030)
Con	−0.021	−0.112
	(0.026)	(0.073)
Con * PR	0.029	0.025
	(0.045)	(0.044)
Con * PC	−0.081*	−0.078*
	(0.032)	(0.032)

Selection equations (Ivs)	Vol	Foc	Con
DD_V, DD_F, DD_C	−0.023***	−0.003	−0.002
	(0.005)	(0.006)	(0.005)
$D1_V, D1_F, D1_C$	0.460***	0.604***	0.323***
	(0.080)	(0.096)	(0.085)
$D2_V, D2_F, D2_C$	0.127*	0.225**	0.144*
	(0.062)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$	0.133*	0.115	0.165**
	(0.057)	(0.077)	(0.061)
$D4_V, D4_F, D4_C$	−0.061	−0.070	0.147*
	(0.053)	(0.064)	(0.067)
$D5_V, D5_F, D5_C$	0.020	−0.025	−0.018
	(0.053)	(0.074)	(0.060)
Error correlations			
$\rho_{VD}, \rho_{FD}, \rho_{CD}$	−0.113***	−0.041	0.070+
	(0.031)	(0.030)	(0.042)
ρ_{VF}, ρ_{VC}		0.462***	−0.407***
		(0.049)	(0.051)
ρ_{FC}			0.135*
			(0.059)

Total effect routine patients		
Vol	−0.168**	−0.013
	(0.054)	(0.070)
Foc	−0.219***	−0.167*
	(0.044)	(0.069)
Con	0.007	−0.087
	(0.043)	(0.083)
Total effect complex patients		
Vol	−0.021	0.124*
	(0.036)	(0.058)
Foc	−0.004	0.046
	(0.027)	(0.055)
Con	−0.103***	−0.190*
	(0.030)	(0.077)
Effect differences		
Δ Vol (PR, PC)	−0.147*	−0.137*
	(0.057)	(0.057)
Δ Foc (PR, PC)	−0.215***	−0.208***
	(0.050)	(0.050)
Δ Con (PR, PC)	0.110*	0.103*
	(0.051)	(0.051)
Observations	265,133	265,133
Segments-in-hospitals	2,067	2,067

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 20 Probit and simultaneous equations models for seven-day mortality: 4 comorbidity threshold

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.076∗	0.066		
	(0.033)	(0.055)		
Vol * PR	−0.032	−0.028		
	(0.036)	(0.036)		
Vol * PC	0.120∗	0.117∗		
	(0.050)	(0.050)		
Foc	−0.040+	0.011		
	(0.023)	(0.053)		
Foc * PR	−0.179∗∗∗	−0.173∗∗∗		
	(0.033)	(0.034)		
Foc * PC	0.035	0.035		
	(0.040)	(0.040)		
Con	−0.081∗∗	−0.175∗		
	(0.026)	(0.073)		
Con * PR	0.112∗∗	0.107∗∗		
	(0.035)	(0.034)		
Con * PC	−0.051	−0.050		
	(0.043)	(0.043)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023∗∗∗	−0.003	−0.002
			(0.005)	(0.006)	(0.006)
$D1_V, D1_F, D1_C$			0.460∗∗∗	0.603∗∗∗	0.324∗∗∗
			(0.080)	(0.096)	(0.085)
$D2_V, D2_F, D2_C$			0.127∗	0.225∗∗	0.144∗
			(0.062)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$			0.134∗	0.115	0.165∗∗
			(0.057)	(0.077)	(0.061)
$D4_V, D4_F, D4_C$			−0.061	−0.069	0.148∗
			(0.053)	(0.063)	(0.067)
$D5_V, D5_F, D5_C$			0.020	−0.024	−0.017
			(0.053)	(0.074)	(0.060)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.112∗∗∗	−0.041	0.072+
			(0.031)	(0.030)	(0.041)
ρ_{VF}, ρ_{VC}C				0.462∗∗∗	−0.407∗∗∗
				(0.049)	(0.051)
ρ_{FC}					0.136∗
					(0.059)
Total effect routine patients					
Vol	−0.109∗∗	0.037			
	(0.038)	(0.059)			
Foc	−0.219∗∗∗	−0.161∗∗			
	(0.029)	(0.059)			
Con	0.030	−0.068			
	(0.031)	(0.075)			
Total effect complex patients					
Vol	0.043	0.182∗			
	(0.054)	(0.071)			
Foc	−0.005	0.046			
	(0.039)	(0.062)			
Con	−0.133∗∗	−0.224∗∗			
	(0.044)	(0.084)			
Effect differences					
Δ Vol (PR, PC)	−0.152∗∗	−0.145∗			
	(0.058)	(0.057)			
Δ Foc (PR, PC)	−0.214∗∗∗	−0.208∗∗∗			
	(0.047)	(0.047)			
Δ Con (PR, PC)	0.163∗∗	0.156∗∗			
	(0.053)	(0.053)			
Observations	265,133	265,133			
Segments-in-hospitals	2,067	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Note that this revised complexity classification is somewhat less balanced than the complexity classification used in the main paper. We now classify 7% of the sample patients as complex, 39% as benchmark and 54% as routine patients as opposed to the original 14% complex, 44% benchmark and 41% routine patients. The estimation results analogous to Table 1 in the main paper are reported in Table 21. In line with the results reported in the main paper, we find partial support for our volume hypothesis and full support for our focus and concentration hypothesis.

9.3. Comorbidities and process uncertainty

We have used process uncertainty as our theoretical lense in the hypothesis development in the main paper. Process uncertainty relates to a lack of information that allows confident service process decisions at the start of the service episode. One can therefore argue that comorbidities that are known at the start of the service episode do not increase process uncertainty. It would therefore be useful to only use those comorbidities that are uncovered later in the service episode to define complex patients - as patients with multiple such comorbidities are more clearly associated with high process uncertainty. Unfortunately, our data does not have information about the time when a comorbidity was discovered. We therefore interviewed five physicians of large general hospitals and asked them to indicate which comorbidities they believe to be largely known upon hospital admission. Specifically, we used a 4-point Likert-scale against each comorbidity, indicating whether the comorbidity is almost always known (=1), predominantly known (=2), predominantly unknown (=3) and almost never known (=4) at the time of admission. The results show variation across comorbidities and across physicians. In fact, none of the comorbidities was rated as always known by all five physicians. For each comorbidity, we computed the average score across all physicians, which we then used to rank the comorbidities. We deleted all comorbidities whose score was lower than the median score (2.6), leaving us with comorbidities that these five physicians regard as most likely to be uncovered later in the service process. These are pulmonary circulation disorders, hypothyroidism, liver disease, lymphoma, solid tumor without metastasis, rheumatoid arthritis/collagen, vascular diseases, coagulopathy, drug abuse, psychoses, and depression. We then call an emergency patient complex if the patient has at least one of these comorbidities (note that one out of 10 is a similar proportion to 3 out of 31 comorbidities used before). The sample was then composed of 115,399 benchmark patients, 130,384 routine patients, and 19,350 complex patients. We replicated the analysis and while the effect difference for volume is only weakly significant, our hypothesis concerning focus and concentration remain fully confirmed (see Table 22). All interviewed physicians are based at large tertiary hospitals and may thus not be representative for general hospitals. However, these hospitals usually serve as the last point of call in the hospital chain and frequently provide care for patients being transferred from other hospitals and

Table 21 Probit and simultaneous equations models for seven-day mortality: highest resource consumption

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.048	0.100+		
	(0.034)	(0.055)		
Vol * PR	−0.095∗	−0.089∗		
	(0.037)	(0.037)		
Vol * PC	0.025	0.023		
	(0.048)	(0.047)		
Foc	−0.054∗	0.004		
	(0.023)	(0.052)		
Foc * PR	−0.142∗∗∗	−0.136∗∗∗		
	(0.034)	(0.034)		
Foc * PC	0.047	0.045		
	(0.044)	(0.044)		
Con	−0.063∗	−0.157∗		
	(0.025)	(0.073)		
Con * PR	0.109∗∗	0.105∗∗		
	(0.034)	(0.034)		
Con * PC	−0.146∗∗∗	−0.140∗∗∗		
	(0.042)	(0.042)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023∗∗∗	−0.003	−0.002
			(0.005)	(0.006)	(0.006)
$D1_V, D1_F, D1_C$			0.460∗∗∗	0.603∗∗∗	0.323∗∗∗
			(0.080)	(0.096)	(0.085)
$D2_V, D2_F, D2_C$			0.128∗	0.225∗∗	0.144∗
			(0.063)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$			0.133∗	0.115	0.165∗∗
			(0.057)	(0.077)	(0.061)
$D4_V, D4_F, D4_C$			−0.061	−0.069	0.147∗
			(0.053)	(0.063)	(0.067)
$D5_V, D5_F, D5_C$			0.020	−0.025	−0.017
			(0.053)	(0.074)	(0.060)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.117∗∗∗	−0.046	0.073+
			(0.031)	(0.029)	(0.042)
ρ_{VF}, ρ_{VC}				0.461∗∗∗	−0.406∗∗∗
				(0.049)	(0.051)
ρ_{FC}					0.136∗
					(0.059)
Total effect routine patients					
Vol	−0.143∗∗∗	0.011			
	(0.039)	(0.059)			
Foc	−0.195∗∗∗	−0.132∗			
	(0.029)	(0.059)			
Con	0.046	−0.052			
	(0.032)	(0.076)			
Total effect complex patients					
Vol	−0.023	0.124+			
	(0.049)	(0.067)			
Foc	−0.007	0.049			
	(0.042)	(0.065)			
Con	−0.209∗∗∗	−0.297∗∗∗			
	(0.047)	(0.083)			
Effect differences					
Δ Vol (PR, PC)	−0.120∗	−0.122∗			
	(0.029)	(0.059)			
Δ Foc (PR, PC)	−0.188∗∗∗	−0.181∗∗∗			
	(0.051)	(0.051)			
Δ Con (PR, PC)	0.255∗∗∗	0.248∗∗∗			
	(0.049)	(0.049)			
Observations	265,133	265,133			
Segments-in-hospitals	2,067	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

are thus more likely to have more information than the patients' first hospital. Thus, if comorbidities are predominantly not known in these tertiary hospitals, we may reasonable assume that this knowledge is also not present in upstream hospitals. This renders our approach conservative.

9.4. Admission status and process uncertainty

In addition to the comorbidity burden, we used admission type to classify complex patients. In doing so, we assume that for patients admitted as an emergency case, less information is available to make informed decisions about the patient's diagnosis and treatment trajectory compared to patients with a planned admission. Besides presenting with less information, emergency patients may, however, also differ from non-emergency patients in terms of the acuity of their disease and therefore the results we find may not be entirely due to process uncertainty but due to the hospital's ability to respond rapidly to acute conditions. While we cannot isolate the acuity aspect from process uncertainty completely, we can exploit more granular data that we have available for some of our sample hospitals, with information about which procedures were applied and when, to analyse these patients' diagnosis and treatment trajectories in more detail, and specifically the differences between routine and complex patients.

9.4.1. Task variety Resolving uncertainty related to diagnosis and treatment involves search processes. We therefore expect the variety of executed tasks to be higher for patients with higher process uncertainty and, if process uncertainty is related to our complexity measure, that this would also be the case for complex patients, relative to routine patients. To test this, we choose groups of related procedures and calculated the number of different procedure groups for every patient (for the aforementioned subsample of hospitals) with the procedure groups based upon the German procedure classification (DIMDI - Deutsches Institut für Medizinische Dokumentation und Information 2014). We find that complex patients have an average of 2.8 (SD: 1.9), benchmark patients an average of 2.6 (SD: 1.8), and routine patients an average of 2.2 (SD: 1.5) types of procedures. We additionally test whether these differences are statistically significant by means of a Poisson regression, with the number of procedure groups as dependent variable, controlling for patient demographics via age and gender, substantial variation between segments via segment fixed effects and for variation across hospitals by means of hospital characteristics (beds, teaching status, ownership). We obtain a negative coefficient for routine patients ($\beta = -0.176, p < 0.001$) and a positive coefficient for complex patients ($\beta = 0.095, p < 0.001$), with a significant difference between the coefficients ($-0.176 - 0.095 = -0.271, p < 0.001$). These results offer some support for our argument that our complexity measure also captures process uncertainty.

Table 22 Probit and simultaneous equations models for seven-day mortality: Unknown comorbidities only

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.054	0.104+		
	(0.034)	(0.056)		
Vol * PR	−0.067+	−0.063+		
	(0.036)	(0.036)		
Vol * PC	0.030	0.032		
	(0.052)	(0.051)		
Foc	−0.059∗	0.005		
	(0.023)	(0.053)		
Foc * PR	−0.127∗∗∗	−0.122∗∗∗		
	(0.033)	(0.033)		
Foc * PC	0.076+	0.079+		
	(0.043)	(0.043)		
Con	−0.071∗∗	−0.169∗		
	(0.027)	(0.074)		
Con * PR	0.084∗	0.080∗		
	(0.034)	(0.034)		
Con * PC	−0.046	−0.048		
	(0.045)	(0.044)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023∗∗∗	−0.003	−0.002
			(0.005)	(0.006)	(0.006)
$D1_V, D1_F, D1_C$			0.460∗∗∗	0.603∗∗∗	0.324∗∗∗
			(0.080)	(0.096)	(0.085)
$D2_V, D2_F, D2_C$			0.127∗	0.224∗∗	0.144∗
			(0.062)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$			0.133∗	0.115	0.165∗∗
			(0.057)	(0.077)	(0.061)
$D4_V, D4_F, D4_C$			−0.061	−0.068	0.147∗
			(0.053)	(0.063)	(0.067)
$D5_V, D5_F, D5_C$			0.020	−0.024	−0.017
			(0.053)	(0.074)	(0.060)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.124∗∗∗	−0.050+	0.076+
			(0.031)	(0.030)	(0.041)
ρ_{VF}, ρ_{VC}				0.462∗∗∗	−0.407∗∗∗
				(0.049)	(0.051)
ρ_{FC}					0.136∗
					(0.060)
Total effect routine patients					
Vol	−0.120∗∗∗	0.041			
	(0.038)	(0.057)			
Foc	−0.185∗∗∗	−0.117∗			
	(0.028)	(0.057)			
Con	0.013	−0.088			
	(0.030)	(0.074)			
Total effect complex patients					
Vol	−0.024	0.136+			
	(0.054)	(0.071)			
Foc	0.017	0.083			
	(0.042)	(0.065)			
Con	−0.117∗∗	−0.217∗∗			
	(0.043)	(0.081)			
Effect differences					
Δ Vol (PR, PC)	−0.096+	−0.095+			
	(0.058)	(0.057)			
Δ Foc (PR, PC)	−0.202∗∗∗	−0.200∗∗∗			
	(0.049)	(0.048)			
Δ Con (PR, PC)	0.130∗∗	0.128∗			
	(0.051)	(0.050)			
Observations	265,133	265,133			
Segments-in-hospitals	2,067	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

9.4.2. Timing of procedures: Date of surgery Resolving uncertainty related to diagnosis and treatment is also time-consuming. Complex patients whose diagnosis and treatment trajectory is unknown in advance pose a high degree of uncertainty and sufficient time is required to decide upon the feasibility and appropriateness of specific procedures. Routine patients, on the other hand, present with less uncertainty and are more likely to have a treatment plan in place upon admission. In this case, less time is required for service planning. One aspect of timing that we can identify in our data is the time of surgery (for patients who had surgery). While the probability of surgery on the admission day (day=0) is similar across patient types (routine, benchmark, complex), the probability of surgery on day 1 is more than twice as high for routine patients compared to complex patients and from day 4 onwards, the probability is higher for complex patients than routine patients (see Figure 6). We find that routine patients that undergo surgery are on average operated on day 2.8 (SD: 4.6), benchmark patients on day 4.3 (SD: 6.3) and complex patients on day 5.9 (SD: 7.6). We additionally test whether the differences are statistically significant by regressing surgery date on patient types, controlling for patient demographics via age and gender, substantial variation between segments via segment fixed effects and for variation across hospitals by means of hospital characteristics (beds, teaching status, ownership). We find a negative coefficient for routine patients ($\beta = -1.251, p < 0.001$) and a positive coefficient for complex patients ($\beta = 1.449, p < 0.001$), with a significant difference between them ($-1.251 - 1.449 = -2.700, p < 0.001$), these results lend additional support to our argument that an increase in complexity is associated with longer search processes.

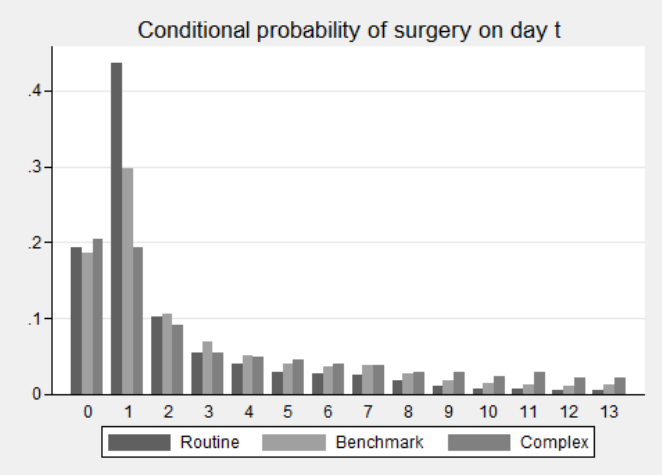


Figure 6 Probability of surgery on day t, conditional on having surgery

10. Heterogeneity in segments: Subsample analyses

10.1. Diseases of the circulatory system

In the main paper we estimate effects across multiple segments simultaneously, in order to maximize statistical power. However, this may lead to concerns about heterogeneity as the segment effects may differ substantially. We therefore conduct analogous estimations for subsamples in this section. In order to maximize statistical power, we first re-estimate the model for all patients in the ICD chapter "Diseases of the circulatory system", which is the largest ICD chapter in our data, accounting for 37% of the patients. This chapter consists of only five patient segments (as defined by ICD blocks). Table 23 shows the estimation results, with the relevant total effect and effect differences supporting all hypotheses. Note that, in contrast to the main paper, this subsample does not show evidence for endogeneity in the volume–mortality relationship (ρ is not significant).

10.2. Six high-risk conditions

Finally, we consider a subsample of six high-risk conditions "for which mortality has been shown to vary substantially across institutions and for which evidence suggests that high mortality may be associated with deficiencies in the quality of care"(Agency for Healthcare Research and Quality 2015): acute myocardial infarction, stroke, congestive heart failure, gastrointestinal hemorrhage, hip replacement after fracture, and pneumonia. Table 24 reports the results for the six high-risk conditions. This subsample does not provide evidence of a volume-selection effect and while the probit model fully supports our focus and concentration hypotheses, the results do not confirm the volume hypothesis at conventional significance levels. Taken together with the results for the subsample of diseases of the circulatory system, this indicates segment-specific heterogeneity. While our hypotheses are (partially) supported across our main segments and thus provide us with an average effect, this does not have to hold for every segment in our main sample.

11. Continuous independent variables

In the main paper, we use dichotomized versions of the continuous variables volume, focus and concentration level, because this approximate heterogeneous nonlinear relationships between these variables and the patient's latent health index better than a linear approximation. While we believe dichotomization to be the most appropriate model specification for our context, it is nevertheless interesting to see whether the results are robust when standardized continuous variables are used. Noite that outliers are particularly problematic when nonlinear relationships are approximated by a linear function. We therefure curtailed our sample to avoid such outlier effects. Specifically, we restrict our sample to segments-in-hospitals where the volume z-score is below 3, the focus z-score is below 3, and the concentration z-score is above -3 (see Figure 7).

Table 23 Probit and simultaneous equations models for seven-day mortality: Diseases of the circulatory system

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.132+	−0.087		
	(0.075)	(0.106)		
Vol * PR	−0.073	−0.070		
	(0.074)	(0.074)		
Vol * PC	0.127+	0.126+		
	(0.066)	(0.066)		
Foc	−0.004	0.012		
	(0.037)	(0.093)		
Foc * PR	−0.215**	−0.214**		
	(0.072)	(0.072)		
Foc * PC	0.009	0.009		
	(0.053)	(0.053)		
Con	−0.085+	−0.076		
	(0.046)	(0.128)		
Con * PR	0.089	0.087		
	(0.065)	(0.065)		
Con * PC	−0.079	−0.078		
	(0.053)	(0.054)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.027***	−0.012	0.017+
			(0.008)	(0.012)	(0.009)
$D1_V, D1_F, D1_C$			0.326*	0.374*	0.451**
			(0.149)	(0.186)	(0.144)
$D2_V, D2_F, D2_C$			0.112	0.232+	0.262*
			(0.106)	(0.128)	(0.117)
$D3_V, D3_F, D3_C$			0.233*	0.067	0.333**
			(0.106)	(0.137)	(0.118)
$D4_V, D4_F, D4_C$			−0.146	−0.191	0.053
			(0.094)	(0.121)	(0.122)
$D5_V, D5_F, D5_C$			0.074	−0.263*	0.042
			(0.095)	(0.134)	(0.103)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.028	−0.015	0.003
			(0.061)	(0.052)	(0.078)
ρ_{VF}, ρ_{VC}				0.417***	−0.626***
				(0.095)	(0.074)
ρ_{FC}					0.116
					(0.111)
Total effect routine patients					
Vol	−0.205*	−0.157			
	(0.083)	(0.116)			
Foc	−0.219***	−0.202+			
	(0.063)	(0.110)			
Con	0.003	0.116			
	(0.068)	(0.144)			
Total effect complex patients					
Vol	0.005	0.039			
	(0.081)	(0.106)			
Foc	−0.005	0.021			
	(0.048)	(0.092)			
Con	−0.164**	−0.154			
	(0.056)	(0.137)			
Effect differences					
Δ Vol (PR, PC)	−0.200*	−0.195*			
	(0.085)	(0.086)			
Δ Foc (PR, PC)	−0.224**	−0.223**			
	(0.075)	(0.075)			
Δ Con (PR, PC)	0.168*	0.166*			
	(0.081)	(0.081)			
Observations	94,622	94,622			
Segments-in-hospitals	296	296			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

Table 24 Probit and simultaneous equations models for seven-day mortality: Six high risk conditions

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.057	−0.055		
	(0.048)	(0.095)		
Vol * PR	0.051	0.051		
	(0.063)	(0.063)		
Vol * PC	0.125∗	0.125∗		
	(0.063)	(0.063)		
Foc	−0.031	0.002		
	(0.033)	(0.082)		
Foc * PR	−0.203∗∗∗	−0.202∗∗∗		
	(0.059)	(0.059)		
Foc * PC	0.008	0.008		
	(0.052)	(0.052)		
Con	−0.063	−0.061		
	(0.042)	(0.119)		
Con * PR	0.133∗	0.132∗		
	(0.060)	(0.060)		
Con * PC	−0.055	−0.055		
	(0.054)	(0.053)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.043∗∗∗	−0.006	0.019∗
			(0.007)	(0.010)	(0.009)
$D1_V, D1_F, D1_C$			0.240+	0.599∗∗∗	0.730∗∗∗
			(0.125)	(0.157)	(0.121)
$D2_V, D2_F, D2_C$			0.219∗	0.250∗	0.169+
			(0.094)	(0.116)	(0.100)
$D3_V, D3_F, D3_C$			0.112	0.089	0.053
			(0.094)	(0.115)	(0.092)
$D4_V, D4_F, D4_C$			−0.135+	−0.091	−0.092
			(0.081)	(0.101)	(0.092)
$D5_V, D5_F, D5_C$			0.010	−0.039	−0.006
			(0.083)	(0.112)	(0.091)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.005	−0.022	−0.003
			(0.052)	(0.049)	(0.065)
ρ_{VF}, ρ_{VC}				0.329∗∗∗	−0.436∗∗∗
				(0.076)	(0.077)
ρ_{FC}					0.182∗
					(0.085)
Total effect routine patients					
Vol	−0.006	−0.004			
	(0.061)	(0.105)			
Foc	−0.234∗∗∗	−0.200∗			
	(0.051)	(0.091)			
Con	0.070	0.071			
	(0.055)	(0.120)			
Total effect complex patients					
Vol	0.068	0.070			
	(0.058)	(0.105)			
Foc	−0.023	0.011			
	(0.043)	(0.084)			
Con	−0.118∗	−0.116			
	(0.047)	(0.117)			
Effect differences					
Δ Vol (PR, PC)	−0.074	−0.074			
	(0.073)	(0.074)			
Δ Foc (PR, PC)	−0.211∗∗	−0.210∗∗			
	(0.068)	(0.068)			
Δ Con (PR, PC)	0.188∗∗	0.187∗∗			
	(0.067)	(0.067)			
Observations	62,470	62,470			
Segments-in-hospitals	575	575			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

We also computed new instrumental variables, based on the continuous variables. For each independent variable (volume, focus, concentration), we use a set of 5 variables D_{ik} ($k \in \{1, \ldots, 5\}$) which are equal to the standardized (volume-, focus-, concentration-) measure of the k-th nearest hospital to patient i. Although the differential distance is a continuous variable, it captures the incremental difference between two hospitals of specific types. However, this requires categorization of hospitals and as such categorization of the continuous measures which isn't aligned with the approach taken here. Therefore, the model does not incorporate this second type of instrumental variable.

The results of the linear model are summarized in Table 25. In contrast to the dichotomous model in Table 1 of the main paper, the correlations ρ_{VD}, ρ_{FD} and ρ_{CD} between the selection equations and the outcome equation (D) are not significant and the model therefore does not identify endogeneity as a critical issue. We do not find the model convincing, though, because the significance pattern of the focus selection equation (second panel, column 3) suggests that the instruments are weak for focus. We therefore refrain from interpreting the results of this model and include it only for completion. We also point out that, if there is no endogeneity, then the probit results should be interpreted instead of the results of the statistically less powerful multivariate probit model. These results (column 1) are in line with the endogeneity controlled results of column 2 in Table 1 in the main paper: Volume does not have a statitsically significant effect on mortality for routine patients, with some evidence that the effect is negative (increases mortality) for complex patients; routine patients benefit from focus and more so than complex patients, and complex patients benefit from segment concentration, and more so than routine patients.

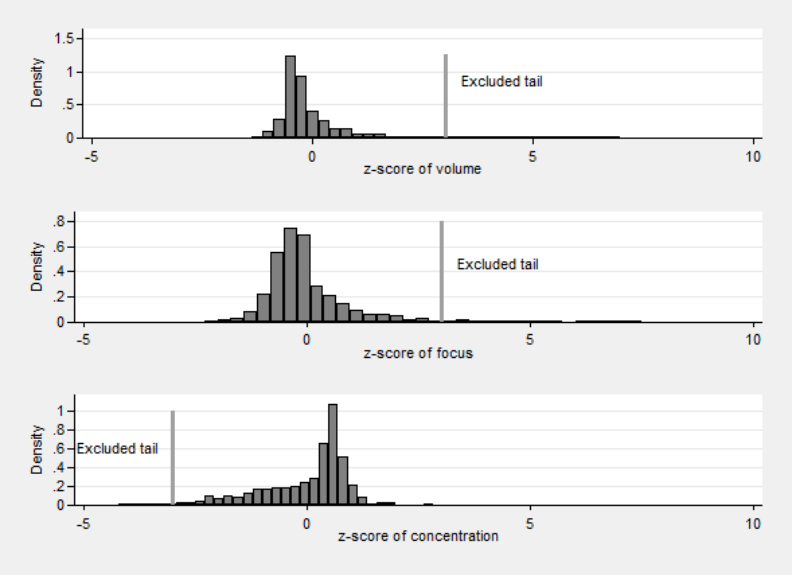


Figure 7 Distribution of standardized continuous variables

Table 25 Probit and simultaneous equations models for standardized independent variables

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.012	0.024		
	(0.029)	(0.069)		
Vol * PR	−0.019	−0.019		
	(0.026)	(0.026)		
Vol * PC	0.070**	0.069**		
	(0.024)	(0.024)		
Foc	−0.086***	−0.085		
	(0.018)	(0.123)		
Foc * PR	−0.084**	−0.084**		
	(0.027)	(0.027)		
Foc * PC	0.018	0.018		
	(0.022)	(0.022)		
Con	−0.012	−0.010		
	(0.014)	(0.053)		
Con * PR	0.065**	0.065**		
	(0.021)	(0.021)		
Con * PC	−0.046*	−0.046*		
	(0.020)	(0.020)		
Selection equation		Vol	Foc	Con
$D1_V, D1_F, D1_C$		0.135***	0.077+	0.249***
		(0.024)	(0.043)	(0.036)
$D2_V, D2_F, D2_C$		0.076***	0.052*	0.154***
		(0.016)	(0.026)	(0.038)
$D3_V, D3_F, D3_C$		0.061***	0.003	0.038
		(0.014)	(0.017)	(0.027)
$D4_V, D4_F, D4_C$		0.011	0.004	0.045
		(0.012)	(0.015)	(0.030)
$D5_V, D5_F, D5_C$		0.036*	−0.014	0.039
		(0.017)	(0.014)	(0.028)
Error correlations				
$\rho_{VD}, \rho_{FD}, \rho_{CD}$		−0.025	−0.010	0.007
		(0.038)	(0.088)	(0.040)
ρ_{VF}, ρ_{VC}			0.341***	−0.343***
			(0.043)	(0.045)
ρ_{FC}				0.066
				(0.055)
Total effect routine patients				
Vol	−0.033	0.005		
	(0.030)	(0.070)		
Foc	−0.171***	−0.169		
	(0.024)	(0.123)		
Con	0.053**	0.055		
	(0.019)	(0.054)		
Total effect complex patients				
Vol	0.057+	0.093		
	(0.030)	(0.068)		
Foc	−0.068**	−0.067		
	(0.022)	(0.126)		
Con	−0.058**	−0.056		
	(0.019)	(0.054)		
Effect differences				
Δ Vol (PR, PC)	−0.089**	−0.088**		
	(0.029)	(0.029)		
Δ Foc (PR, PC)	−0.103***	−0.102***		
	(0.019)	(0.029)		
Δ Con (PR, PC)	0.111***	0.111***		
	(0.025)	(0.025)		
Observations	197,921	197,921		
Segments-in-hospitals	1,970	1,970		

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10

12. Alternative model specifications

12.1. Standard errors clustered at the hospital level

n our main model, we clustered standard errors at the segment-in-hospital level. This is in line with the assumption that patients, who receive service in one disease segment (such as ischemic heart diseases) are independent from patients, who receive service in another disease segment (such as malignant neoplasm of digestive organs) despite being admitted to the same hospital. However, even though the degree of dependence is arguably higher within a hospital-segment than across hospital-segments, we cannot neglect residual dependency at the hospital level. Therefore, we test the robustness of our results and cluster standard errors at the hospital level (Table 26). The significance levels concerning the total effects and effect differences between routine and complex patients remain in line with the results reported in the paper, only for the differential volume-effect it is slightly weaker ($p=0.058$ instead of $p<0.05$).

12.2. Mixed effect probit model

Our econometric models account for the hierarchy in our data by means of cluster-specific probit models with an unobserved cluster-specific effect ν_{sh}. Given that our independent variables do not vary at the segment-within-hospital level, we cannot estimate the cluster-specific effect as fixed effect but rather treated it as random and integrated it out. As an alternative, we also estimated a model that incorporated random effects explicitly, as a standard random effects equation, and estimated probit models with random effects at the segment-in-hospital level, using the meprobit command in STATA 14 and endogeneity controlled models with random effects by means of the user written STATA command cmp (Roodman 2011). Table 27 provides the results. The results concerning the total effects and effect differences between routine and complex patients remain in line with the results reported in the paper.

12.3. Linear probability models

We estimated linear probability models as alternatives to probit models. However, these models were a poor fit to our data and predicted a negative mortality probability for more than 20% of the data. It is well known that linear probability models are severely biased and inconsistent when they predict a substantial proportion of probabilities outside the unit interval and that the bias grows with the proportion of predictions that fall outside this interval (Horrace and Oaxaca 2006). The linear probability specification is therefore inappropriate for our sample and we refrain from reporting the results.

12.4. Survival models with discharge as a competing risk

Since we consider mortality during an observation period of seven days, we may be concerned about two types of censoring: Patients may die after the first seven days or patients may die outside of

Table 26 Probit and simultaneous equations models: SE clustered at the hospital level

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.106**	0.035		
	(0.036)	(0.064)		
Vol * PR	0.011	0.016		
	(0.046)	(0.046)		
Vol * PC	0.129**	0.127**		
	(0.043)	(0.042)		
Foc	−0.057*	−0.007		
	(0.024)	(0.051)		
Foc * PR	−0.196***	−0.189***		
	(0.029)	(0.029)		
Foc * PC	0.038	0.038		
	(0.030)	(0.029)		
Con	−0.056*	−0.144+		
	(0.023)	(0.080)		
Con * PR	0.094**	0.089**		
	(0.033)	(0.089)		
Con * PC	−0.071*	−0.068*		
	(0.032)	(0.032)		
Selection equations (IVs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023***	−0.003	−0.002
			(0.007)	(0.009)	(0.005)
$D1_V, D1_F, D1_C$			0.460***	0.603***	0.323***
			(0.123)	(0.160)	(0.090)
$D2_V, D2_F, D2_C$			0.127	0.225**	0.144
			(0.102)	(0.076)	(0.090)
$D3_V, D3_F, D3_C$			0.134*	0.115	0.165**
			(0.066)	(0.076)	(0.062)
$D4_V, D4_F, D4_C$			−0.061	−0.069	0.147*
			(0.062)	(0.067)	(0.074)
$D5_V, D5_F, D5_C$			0.020	−0.024	−0.018
			(0.071)	(0.072)	(0.065)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.109**	−0.041	0.068+
			(0.039)	(0.029)	(0.051)
ρ_{VF}, ρ_{VC}				0.462***	−0.407***
				(0.075)	(0.062)
ρ_{FC}					0.136*
					(0.060)
Total effect routine patients					
Vol	−0.094+	0.051			
	(0.054)	(0.075)			
Foc	−0.253***	−0.196***			
	(0.031)	(0.057)			
Con	0.038	−0.055			
	(0.029)	(0.087)			
Total effect complex patients					
Vol	0.024	0.162*			
	(0.047)	(0.077)			
Foc	−0.018	0.032			
	(0.031)	(0.058)			
Con	−0.127***	−0.212*			
	(0.033)	(0.092)			
Effect differences					
Δ Vol (PR, PC)	−0.118*	−0.111+			
	(0.059)	(0.058)			
Δ Foc (PR, PC)	−0.234***	−0.228***			
	(0.037)	(0.037)			
Δ Con (PR, PC)	0.165***	0.158***			
	(0.041)	(0.042)			
Observations	265,133	265,133			
Hospitals	60	60			

Standard errors clustered on hospitals; controls included as per Table 2.***
p<0.001, ** p<0.01, * p<0.05 + p<0.10

Table 27 Mixed-effect probit and simultaneous equations models for seven-day mortality: Hospital-Segment RE

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.111***	−0.007		
	(0.033)	(0.051)		
Vol * PR	0.012	0.015		
	(0.037)	(0.047)		
Vol * PC	0.132***	0.130**		
	(0.039)	(0.050)		
Foc	−0.066**	−0.022		
	(0.023)	(0.050)		
Foc * PR	−0.187***	−0.183***		
	(0.031)	(0.043)		
Foc * PC	0.036	0.036		
	(0.031)	(0.044)		
Con	−0.047*	−0.090		
	(0.024)	(0.058)		
Con * PR	0.092**	0.090*		
	(0.029)	(0.041)		
Con * PC	−0.071*	−0.069		
	(0.033)	(0.046)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023***	−0.003***	−0.002***
			(0.001)	(0.001)	(0.001)
$D1_V, D1_F, D1_C$			0.460***	0.604***	0.323***
			(0.009)	(0.007)	(0.007)
$D2_V, D2_F, D2_C$			0.127***	0.226***	0.144***
			(0.008)	(0.006)	(0.006)
$D3_V, D3_F, D3_C$			0.134***	0.116***	0.165***
			(0.007)	(0.006)	(0.006)
$D4_V, D4_F, D4_C$			−0.061***	−0.070***	0.147***
			(0.007)	(0.006)	(0.006)
$D5_V, D5_F, D5_C$			0.020**	−0.024***	−0.017**
			(0.007)	(0.006)	(0.006)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.077**	−0.035	0.036
			(0.025)	(0.026)	(0.032)
ρ_{VF}, ρ_{VC}				0.462***	−0.406***
				(0.004)	(0.004)
ρ_{FC}					0.136***
					(0.004)
Total effect routine patients					
Vol	−0.099*	0.009			
	(0.040)	(0.059)			
Foc	−0.253***	−0.206***			
	(0.030)	(0.058)			
Con	0.045	0.000			
	(0.028)	(0.062)			
Total effect complex patients					
Vol	0.021	0.124*			
	(0.041)	(0.060)			
Foc	−0.030	0.013			
	(0.031)	(0.056)			
Con	−0.118***	−0.159*			
	(0.032)	(0.066)			
Effect differences					
Δ Vol (PR, PC)	−0.120**	−0.115*			
	(0.046)	(0.058)			
Δ Foc (PR, PC)	−0.223***	−0.219***			
	(0.037)	(0.052)			
Δ Con (PR, PC)	0.163***	0.160**			
	(0.037)	(0.051)			
Random effects					
σ^2, σ	0.027***	0.164***			
	(0.003)	(0.010)			
Observations	265,133	265,133			
Groups	2,670	2,670			

Standard errors in parentheses; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10.

the hospital if they are discharged prior to day seven. While death after seven days constitutes non-informative censoring and does not affect the seven-day mortality estimates, discharge prior to Day 7 is informative because it is affected by the patient's health status. We therefore replicate our analysis with discharge as a competing risk within a discrete survival analysis model. In line with Kuntz et al. (2015), we keep records of patients discharged prior to day t=7 in the data set. This assumes that all patients discharged prior to day 7 would have survived had they stayed within the hospital. As such, it renders our seven-day mortality estimation approach conservative (see Kuntz et al. (2015) for details). Table 28 provides the results of the survival analysis. In line with the main paper, we find partial effect for the volume-hypothesis - i.e. a significant effect difference between routine and complex patient yet not a beneficial total effect for routine patients -, and full support for our focus-hypothesis. Opposed to the main paper, the concentration-hypothesis is no longer fully supported, the total effect for complex patients is no longer significant. However, the effect difference between routine and complex patients still remains significant lending partial support to our concentration-hypothesis.

12.5. Disentangling emergency and comorbidity effects

As previously outlined, we use admission type and the patient's comorbidity burden to classify different levels of patient complexity. One may therefore ask the question which of the two factors is the dominant moderator of the independent variables. Is patient complexity more likely to be caused by the admission type or by the comorbidity burden or by the combination of both? In the main paper, our benchmark patients were either elective patients with a high comorbidity burden or emergency patients with a low comorbidity burden. We now split the benchmark patients into two categories: Benchmark I, containing high-comorbidity electives (BCom), and Benchmark II capturing low-comorbidity emergencies (BEm). Pairwise comparison of the patient types then allows us to disentangle the admission type from the comorbidity effect. The results are summarized in Tables 29 and 30. The "Total effects..." panels in the second half of Table 29 show that the volume effect remains insignificant for benchmark patients, independently of emergency (BEm) or elective (BCom) types. As in Table 1 of the main paper, only complex patients (PC) show a significant volume effect, with higher volume being associated with higher mortality. The focus effect remains significant only for routine patients and the difference in the focus effects between routine and non-routine patient remains highly significant (see Table 30, "Focus" panel, "PR" row). Concentration, however, now has a significant beneficial effect on the emergency patients amongst the benchmark patients. This is entirely in line with our reasoning that concentration of patient segments into single departments is particularly beneficial when there is less information available for these patients, which is particularly relevant for emergency patients. While the coefficient for

Table 28 Probit coefficient of death on day t, conditional on survival up to the end of day t-1

Mortality equation	Probit	Simultaneous equations model
Vol	−0.075∗∗	0.030
	(0.025)	(0.041)
Vol * PR	0.004	0.008
	(0.031)	(0.031)
Vol * PC	0.100∗∗∗	0.099∗∗∗
	(0.030)	(0.030)
Foc	−0.042∗	0.026
	(0.017)	(0.040)
Foc * PR	−0.153∗∗∗	−0.146∗∗∗
	(0.029)	(0.029)
Foc * PC	0.023	0.021
	(0.025)	(0.025)
Con	−0.043∗	−0.025
	(0.019)	(0.052)
Con * PR	0.075∗	0.072∗
	(0.030)	(0.030)
Con * PC	−0.054∗	−0.051∗
	(0.025)	(0.025)
Day 1	0.142∗∗∗	0.142∗∗∗
	(0.016)	(0.016)
Day 2	0.026	0.025
	(0.020)	(0.020)
Day 3	−0.021	−0.021
	(0.020)	(0.020)
Day 4	−0.082∗∗∗	−0.082∗∗∗
	(0.021)	(0.021)
Day 5	−0.098∗∗∗	−0.098∗∗∗
	(0.021)	(0.021)
Day 6	−0.154∗∗∗	−0.153∗∗∗
	(0.022)	(0.022)
Day 7	−0.183∗∗∗	−0.183∗∗∗
	(0.023)	(0.022)

Selection equations (Ivs)	Vol	Foc	Con
DD_V, DD_F, DD_C	−0.023∗∗∗	−0.003	−0.002
	(0.005)	(0.006)	(0.005)
$D1_V, D1_F, D1_C$	0.458∗∗∗	0.604∗∗∗	0.320∗∗∗
	(0.080)	(0.096)	(0.085)
$D2_V, D2_F, D2_C$	0.127∗	0.224∗∗	0.142∗
	(0.063)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$	0.133∗	0.114	0.165∗∗
	(0.057)	(0.077)	(0.061)
$D4_V, D4_F, D4_C$	−0.060	−0.070	0.148∗
	(0.053)	(0.064)	(0.068)
$D5_V, D5_F, D5_C$	0.019	−0.027	−0.017
	(0.053)	(0.074)	(0.060)
Error correlations			
$\rho_{VD}, \rho_{FD}, \rho_{CD}$	−0.075∗∗∗	−0.054∗	−0.001
	(0.022)	(0.022)	(0.030)
ρ_{VF}, ρ_{VC}		0.460∗∗∗	−0.407∗∗∗
		(0.049)	(0.051)
ρ_{FC}			0.136∗
			(0.059)

Total effect routine patients		
Vol	−0.071∗	0.038
	(0.032)	(0.045)
Foc	−0.195∗∗∗	−0.120∗∗
	(0.025)	(0.046)
Con	0.032	0.475
	(0.027)	(0.056)
Total effect complex patients		
Vol	0.025	0.129∗∗
	(0.031)	(0.047)
Foc	−0.020	0.047
	(0.023)	(0.042)
Con	−0.096∗∗∗	−0.076
	(0.025)	(0.057)
Effect differences		
Δ Vol (PR, PC)	−0.098∗	−0.091∗
	(0.039)	(0.038)
Δ Foc (PR, PC)	−0.176∗∗∗	−0.167∗∗∗
	(0.033)	(0.033)
Δ Con (PR, PC)	0.129∗∗∗	0.123∗∗∗
	(0.036)	(0.036)
Observations	2,088,469	2,088,469
Cluster	2,067	2,067

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10

complex patients is larger in magnitude than the coefficient for emergency benchmark patients, Table 30 shows that this difference is not statistically significant ("Concentration" panel, "BEm" row, coeff = 0.020).

Table 29 Probit and simultaneous equations model: splitting benchmark patients

Mortality equation	Probit	Simultaneous equations model		
Vol	−0.127**	0.010		
	(0.043)	(0.063)		
Vol * PR	0.038	0.043		
	(0.050)	(0.050)		
Vol * PC	0.151**	0.149**		
	(0.048)	(0.048)		
Vol * BEm	0.035	0.036		
	(0.049)	(0.048)		
Foc	−0.077*	−0.022		
	(0.033)	(0.058)		
Foc * PR	−0.184***	−0.179***		
	(0.045)	(0.045)		
Foc * PC	0.058	0.056		
	(0.042)	(0.041)		
Foc * BEm	0.050	0.047		
	(0.042)	(0.041)		
Con	−0.014	−0.115		
	(0.036)	(0.077)		
Con * PR	0.048	0.046		
	(0.047)	(0.047)		
Con * PC	−0.105*	−0.101*		
	(0.043)	(0.043)		
Con * BEm	−0.084+	−0.081+		
	(0.045)	(0.045)		
Selection equations (Ivs)			Vol	Foc	Con
DD_V, DD_F, DD_C			−0.023***	−0.003	−0.002
			(0.005)	(0.006)	(0.006)
$D1_V, D1_F, D1_C$			0.460***	0.602***	0.324***
			(0.080)	(0.096)	(0.085)
$D2_V, D2_F, D2_C$			0.127*	0.226**	0.144*
			(0.063)	(0.074)	(0.064)
$D3_V, D3_F, D3_C$			0.133*	0.114	0.165**
			(0.057)	(0.077)	(0.061)
$D4_V, D4_F, D4_C$			−0.061	−0.070	0.147*
			(0.053)	(0.063)	(0.068)
$D5_V, D5_F, D5_C$			0.020	−0.024	−0.018
			(0.053)	(0.074)	(0.060)
Error correlations					
$\rho_{VD}, \rho_{FD}, \rho_{CD}$			−0.109***	−0.042	0.075+
			(0.031)	(0.029)	(0.042)
ρ_{VF}, ρ_{VC}				0.462***	−0.406***
				(0.049)	(0.051)
ρ_{FC}					0.136*
					(0.059)
Total effect low-comorbity electives (PR)					
Vol	−0.089*	0.053			
	(0.043)	(0.061)			
Foc	−0.261***	−0.200***			
	(0.033)	(0.062)			
Con	0.034	−0.070			
	(0.036)	(0.078)			
Total effect high-comorbidity electives (BCom)					
Vol	−0.127**	0.010			
	(0.043)	(0.063)			
Foc	−0.077*	−0.022			
	(0.033)	(0.058)			
Con	−0.014	−0.115			
	(0.036)	(0.077)			
Total effect low-comorbidity emergencies (BEm)					
Vol	−0.092*	0.046			
	(0.040)	(0.058)			
Foc	−0.027	0.025			
	(0.029)	(0.055)			
Con	−0.099**	−0.196*			
	(0.033)	(0.076)			
Total effect high-comorbidity emergencies (PC)					
Vol	0.025	0.160*			
	(0.043)	(0.062)			
Foc	−0.019	0.034			
	(0.031)	(0.058)			
Con	−0.119***	−0.216**			
	(0.035)	(0.079)			
Observations	265,133	265,133			
Cluster	2,067	2,067			

Standard errors clustered on segments-in-hospitals; controls included as per Table 2.*** p<0.001, ** p<0.01, * p<0.05 + p<0.10

Table 30 Effect differences between patient types based on simultaneous equations model in Table 29

	Volume		
	BCom	BEm	PC
PR	0.043	0.007	−0.106∗
	(0.050)	(0.046)	(0.051)
BCom		−0.036	−0.149∗∗
		(0.048)	(0.048)
BEm			−0.114∗
			(0.046)
	Focus		
	BCom	BEm	PC
PR	−0.179∗∗∗	−0.225∗∗∗	−0.234∗∗∗
	(0.045)	(0.041)	(0.043)
BCom		−0.047	−0.056
		(0.041)	(0.041)
BEm			−0.009
			(0.037)
	Concentration		
	BCom	BEm	PC
PR	0.046	0.127∗∗	0.147∗∗
	(0.047)	(0.047)	(0.049)
BCom		0.081+	0.101∗
		(0.044)	(0.043)
BEm			0.020
			(0.040)

*** p<0.001, ** p<0.01, * p<0.05 + p<0.10

References

Agency for Healthcare Research and Quality. 2015. Inpatient safety indicators. URL `http://hcupnet.ahrq.gov/IQIV21R3.htm`. Last access: 2015-11-03.

Clark, JR. 2012. Comorbidity and the limitations of volume and focus as organizing principles. *Medical Care Research and Review* **69**(1) 83–102.

DIMDI - Deutsches Institut für Medizinische Dokumentation und Information. 2014. Ops. URL `http://www.dimdi.de/static/de/klassi/ops/index.htm`.

Gemeinsamer Bundesausschuss. 2016. Qualitätsberichte der Krankenhäuser. URL `https://www.g-ba.de/institution/themenschwerpunkte/qualitaetssicherung/qualitaetsbericht/?`

Horrace, WC, RL Oaxaca. 2006. Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters* **90** 321–327.

KC, DS, C Terwiesch. 2011. The effects of focus on performance: Evidence from California hospitals. *Management Science* **57**(11) 1897–1912.

Kuntz, L, R Mennicken, S Scholtes. 2015. Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science* **61**(4) 754–771.

McClellan, M, BJ McNeil, JP Newhouse. 1994. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *Journal of the American Medical Association* **272**(11) 859–866.

Picard, R. 2012. GEONEAR: Stata module to find nearest neighbors using geodetic distances. *Statistical Software Components* .

Roodman, D. 2011. Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal* **11**(2) 159–206.