Selective referral or learning by doing? An analysis of hospital volume-outcome relationship of vascular procedures

Jakub Červený*

September 13, 2022

Abstract

This paper analyzes the effects of hospital volume on outcomes of patients undergoing percutaneous transluminal angioplasty (PTA) with stent implant in Slovakia between 2014-2019. The volume-outcome relationship is estimated jointly using a discrete factor approach, where choice of hospital is correlated with durations until readmission or death, accounting for observed and unobserved characteristics. The results reveal the importance of controlling for between-hospital differences and selectivity in patient referral. Estimates without hospital fixed effects overstate the positive effect of volume on outcomes, but the results remain statistically significant. Once selectivity is accounted for in the joint correlated model, the positive volume-outcome relationship is not different from zero. Overall, the main driver of the volume-outcome relationship for PTA procedures appears to be related to selective referral and differences in quality of health care providers.

Keywords: Hospital volume-outcome, hospital choice, duration models
JEL-codes: C41, I10

*Institute for Health Care Analyses, Ministry of Health of the Slovak Republic; jakub.cerveny@health.gov.sk

The author would like to thank Noemi Kiss, Ines Stelzer, Martin Mikloš, Martin Šuster, Richard Kališ and Jan van Ours for excellent comments on previous versions of the paper.

1 Introduction

The hospital volume-outcome relationship is a relatively well-researched topic in the health economics literature, with most studies finding that an increase in the volume of procedures is associated with better outcomes for patients. The outcomes are usually measured as mortality, hospital length of stay (LOS) or readmission. The explanation of this observed positive relationship is that "practice makes perfect" or "learning by doing" – hospitals with a higher volume of procedures improve over time thanks to an accumulated knowledge, skills and repetitive routine (Luft et al. (1987)). These findings, combined with increasing returns to scale (i.e. average costs fall as volume increases) prompted policymakers in many countries to reorganize health care systems by imposing minimum procedural volume standards for providers. This policy is known as evidence-based hospital referral (EBHR) limits.[1] Similar policies for the centralization of procedures into specialized hospitals and the introduction of volume limits are currently being proposed as part of health care reform in Slovakia.

From an economic perspective, the rationale behind these proposals are clear – if increased volume of procedures actually leads to better outcomes for patients, the setting of minimum standards and closure of providers not meeting them should lead to an increase in welfare and decrease in costs. If, however, the perceived volume-outcome relationship reflects other mechanisms such as quality differences between hospitals, the closure of providers may restrict access to health care. Although the positive relationship between volume and outcomes is well established, relatively few studies accounted for potential endogeneity or reverse causality in their analyses.

There are several caveats associated with a naive comparison of low-volume hospitals with high-volume hospitals over time. If an econometrician simply regresses the volume of hospital procedures on an outcome indicator, the coefficient capturing the volume effects is likely to be biased due to several issues. First, it is conceivable that perceived quality of health care providers is known among

[1] EBHR limits are a quality metric based primarily on hospital procedural volume. See `https://ratings.leapfroggroup.org/measure/hospital/complex-adult-and-pediatric-surgery`.

patients or doctors referring them from primary care. This leads to a "selective referral" (Luft (1980)), where the volume of procedures becomes inflated, masking the information about quality. Second, the observed differences between hospitals may be explained by cohort effects, also referred to as a case mix. Therefore, using an inadequate set of observed characteristics capturing patient's health status, comorbid conditions as well as ignoring unobserved heterogeneity may lead to biased estimates and incorrect causal conclusions. The volume-outcome relationship could also be explained by self-selection (i.e. surgeons or hospitals more confident or more skilled in the procedures will take on a larger volume of the procedures). Furthermore, large-volume hospitals may seek more skilled or experienced surgeons, while more experienced surgeons may seek busier hospitals.

The main contribution of this paper to the volume-outcome literature is the econometric framework, in which hospital choice is jointly estimated with the effects of volume on patient outcomes to address the issue with selectivity in patient referral. First, a conditional logit model is specified, in which patients face choice of a provider for percutaneous transluminal angioplasty[2] (PTA) with a stent implant. The outcomes of the procedure, measured as 30-day mortality and readmission are then analyzed using similar survival models as used by Hamilton and Hamilton (1997) and jointly estimated with hospital choice using a discrete mixture of unobserved heterogeneity affecting both processes. The joint discrete factor approach was previously used to account for selectivity in labor market studies analyzing effects of sanctions on job search and duration of unemployment (van den Berg et al. (2004), van der Klaauw and van Ours (2013)), and in health economics to analyze effects of drug use on health (van Ours and Williams (2012)).

I use a detailed administrative dataset on all PTA procedures in Slovakia between 2014-2019, which includes detailed information about patient characteristics, comorbid conditions and the type of procedure. Compared to several previous studies on the volume-outcome relationship which mostly focus on in-hospital sur-

[2]PTA is a minimally-invasive vascular procedure, in which a balloon catheter is inserted into a narrowed artery or vessel to improve blood flow. To ensure that the vessel remains open, a stent (either a plastic or a metal tube) is sometimes inserted. PTA with stent implant is most ofted used to threat atherosclerosis (clogging of arteries due to build-up of plaque). Untreated atherosclerosis can lead to coronary artery disease, stroke, myocardial infarction or kidney problems.

vival or length of stay, the dataset used in this paper allows for the tracking of patients after hospital discharge. This could alleviate concerns of a possible bias due to discharge patterns, where high volume hospitals would transfer patients shortly before they die or discharge them early. As will be shown in the following sections, the risk of death or readmission to hospital is highest within the first days after a procedure, therefore focusing solely on in-hospital outcomes might not reveal the full picture.

The results reveal that ignoring selectivity overstates the effect of hospital volume on both 30-day mortality and readmissions. In single-equation models, the estimated volume-outcome effect suggests that a 10% increase in volume of procedures decreases mortality by almost 2.4%. Once differences between hospitals are accounted for, the effect decreases to 1.4%. In the full model, where hospital choice and outcomes are estimated jointly, the effect further drops to 0.9% and is no longer statistically significant. A similar story unfolds for 30-day readmissions, where the volume effect decreases from 1.7% to 0.5% and becomes indistinguishable from zero. This is an important result from the policy perspective, where regionalization of health care and imposition of minimum volume standards may not always achieve the desired outcomes of improved quality. The remainder of the paper is organized as follows: section two summarizes the volume-outcome literature, section three describes the dataset and provides summary statistics, section four describes the econometric framework, section five presents results while section six provides concluding remarks.

2 Review of literature

The results reported in the literature vary substantially and are mostly dependent on the type of procedure as well as on statistical techniques used to estimate the effects. One of the first publications analyzing volume-outcome relationship is a work by Luft et al. (1979), who analyzed 12 types of surgical procedures from more than 1500 hospitals in the United States between 1974-1975. Their findings suggest that mortality of patients undergoing open heart surgery, coronary artery bypass, vascular surgery, transurethral resection of prostate declined by 25 to 41 percent with an increase in volume of procedures. They also report that these effects flatten out at a volume of 10 to 50 procedures per year. No significant effects were observed for vagotomy and cholecystectomy. One of the shortcomings of the analysis is that it used relatively simple, descriptive methods to analyze the relationship between volume and mortality. In a follow-up paper, Luft (1980) extends the analysis and examines the relationship using regression techniques to control for potential confounders. To explore the direction of causality, the authors also estimate a model where the volume and outcome equations are modelled simultaneously. The findings in the single-equation models confirm a positive volume-outcome relationship for procedures as open heart surgery or vascular surgery. However, once the relationship is estimated simultaneously, most of the statistically significant effects disappear, supporting the fact that part of the observed correlation between surgical volume and outcomes can be explained by selective referrals. The authors also note the importance of other factors omitted in their analysis, such as hospital or surgeon fixed effects, which are later explored by other papers and following publications.

Hannan et al. (1992) in their analysis exploit longitudinal data from New York covering a period of six years, allowing for comparison between hospitals as well as within hospitals over time. The effects on in-hospital mortality of patients undergoing surgeries of ruptured and non-ruptured aneurysm are first analyzed using a logistic regression, controlling for hospital and surgeon volume and other observed characteristics. The results suggest a significantly positive effect of higher surgeon volume on in-hospital mortality for ruptured aneurysms, and a positive effect of hospital volume on unruptured aneurysms. To test for a selective referral hypothesis, the authors then track and compare changes in volume of low and high-mortality surgeons over time. They conclude that there is some evidence of an increased volume of surgeries per year for low-mortality surgeons who increased their volume by 43.6 percent, compared to 14 percent for high-mortality surgeons.

Farley and Ozminkowski (1992) analyze volume-outcome relationships using a dataset of almost 500 community hospitals in the USA over eight years. Their econometric framework uses instrumental variables (IV) to control for potential selectivity in referral of patients and exploits the panel structure of the data using hospital fixed effects to control for hospital-specific unobservables. Results suggest that a higher volume leads to lower in-hospital mortality for procedures involving acute myocardial infarction, hernia repair and neonatal respiratory distress syndrome. The positive effect was also observed for coronary artery bypass grafts, however the authors conclude that the result is mostly caused by referral patterns. No significant effects were observed for patients undergoing hip replacement surgery. Similar results reports Hamilton and Hamilton (1997) in their analysis of hip replacement surgeries in the Canadian province of Quebec between 1991-1993. They analyze the effects of hospital volume on LOS using competing-risk duration models with a discrete mixture of unobserved heterogeneity. The authors first present estimates without hospital fixed effects, which suggest a highly significant and positive effect on LOS. However, once the hospital fixed-effects are included, the relationship disappears, suggesting that the effect rather reflects quality differences between hospitals.

Phillips et al. (1995) estimate the effects of hospital volume on outcomes of patients undergoing PTA procedure, focusing on in-hospital mortality, emergency bypass surgery after PTA (indicative of post-procedural complications), LOS and total charges of hospital stay. Their sample includes cross-sectional data from 110 hospitals in California and more than 25,000 patients undergoing the procedure in 1989. Results show that low-volume hospitals had significantly higher rate of adverse events, whereas high-volume hospitals significantly lower. Similar associations are reported for LOS and charges.

Somewhat different results for coronary artery bypass surgery are reported by Gaynor et al. (2005). The authors also employ IV techniques to estimate the effects of volume on in-hospital mortality on a dataset of more than 360,000 surgeries in California between 1983-1999. Contrary to Farley and Ozminkowski (1992), their results suggest a positive effect. The study also explores whether the volume-outcome effect can be attributed to “learning by doing” or to economies of scale, suggesting it is rather the latter. IV regressions are also employed by Tsai et al. (2006), who analyze the relationship on a dataset of patients older than 65 years hospitalized in Ohio between 1991-1997 due to congestive heart failure (CHF). By comparing the simple regressions with IV estimates, their findings reveal that ignoring the selective referral leads to biased results. Once the selectivity was accounted for, no significant effects of higher hospital volume on survival of patients with CHF was observed. The patterns suggestive of selective referral observed by Hannan et al. (1992) are also reported by Barker et al. (2011), who compared specialized and general hospitals in the USA in their analysis of more than 330,000 cardiac revascularizations between 2000 and 2001. This paper also finds evidence that hospitals with lower mortality rates were able to attract more patients. Employing IV regressions to estimate the effects of volume on mortality, their findings reveal that specialized hospitals did not perform better than general hospitals.

A different approach is used by Avdic et al. (2019), who analyze a policy change in budget restrictions of Swedish municipalities. Following a passage of federal law, which effectively forbid local governments to run budget deficits, several cancer clinics had to be closed to cut costs. Closures of providers in affected regions generated sharp changes in the number of cancer surgeries performed at other remaining clinics in the same region, while creating a natural control group in regions in which no clinics were closed. The authors exploit this quasi-experimental variation in hospital volume as instruments in their empirical framework. The results suggests that for complex cancer surgeries involving full/partial mastectomy, colectomy and prostatectomy, the clinics increasing their volume from 70 to 130 procedures per year had lower mortality by approximately 2.7 percentage points. In addition, by examining cumulative volume over several years, subsequent cancer surgeries and complications during surgery, they also find evidence for "learning by doing" effects.

3 Data and descriptive statistics

The analysis is based on administrative data from the National Health Information Center of Slovakia (NHIC). NHIC administers several national health registries, including a claims database on all health expenditure reimbursements. The dataset used in the analysis combines patient-level data on all procedures provided by public health insurance. The databases are linkable through a unique patient identifier obtained at birth. Public health insurance is in general mandatory for every individual who has permanent residency in the Slovak Republic, thus the databases are effectively covering the whole population.

All patients admitted to hospitals who underwent a non-emergency procedure of PTA with stent implant between 1st July 2013 and 31st December 2019 are included in the sample. This includes PTA of carotid and vertebral arteries, peripheral arteries of the extremities and pelvic arteries, renal arteries, celiac or mesenteric artery and arch of the aorta. For each patient, the date of procedure, hospital performing the procedure and other individual characteristics such as age, gender and residence are included. Information on comorbidities was extracted from other registries including primary care procedures and pharmaceutical prescriptions following Bannay et al. (2016) in order to construct a Charlson comorbidity index according to an algorithm for administrative data developed by Quan et al. (2005). PTA is a minimally invasive procedure, where most patients are monitored overnight after the procedure and discharged on next day if there are no complications. The main outcome evaluated in the analysis is 30-day mortality following the PTA procedure. The mortality data are linked to the procedure data from the central death registry. The second outcome considered in the analysis is all-cause 30-day readmission, defined as any readmission to an acute care hospital occurring at least a day or later following discharge[3] after the PTA procedure. Volume of procedures V in each hospital is measured as a number of procedures prior to the current patient, similarly as in Hamilton and Hamilton (1997). More precisely, it is equal to the number of procedures performed in hospital h within the 12-month period prior to the date τ of the current patient's procedure. Therefore, procedures occurring before 1st July 2014 are excluded from the analysis, since it is not possible to determine the hospital volume in the preceding 12 months.

The data related to hospital choice are collected from a ranking maintained by Institute for Social and Economic Reforms (INEKO).[4] Each year since 2014, INEKO publishes quality indicators and rankings as a part of a patient-oriented project for comparison of hospital and health care providers.[5] For each health

[3]Centers for Medicaid and Medicare provide a similar definition of 30-day readmission rate. See `https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2015-ACR-MIF.pdf`

[4]`https://www.ineko.sk/about/about-us`

[5]`https://kdesaliecit.sk` The name of the website translates as "where to get treated?".

care provider, several statistics such as readmission rate, number of physicians per patient, average LOS or total number of patients per year are depicted in the profile. Furthermore, patient satisfaction with the services provided, collected by insurance companies using a standardized questionnaire is also reported in the rankings.[6]

3.1 Descriptive statistics

The following section provides summary statistics and simple comparisons of hospitals included in the sample. In total, 19 hospitals performed PTA stent implants during the sample period, with an average of almost 194 procedures per calendar year. The distribution of hospital volume is depicted in figure 1.

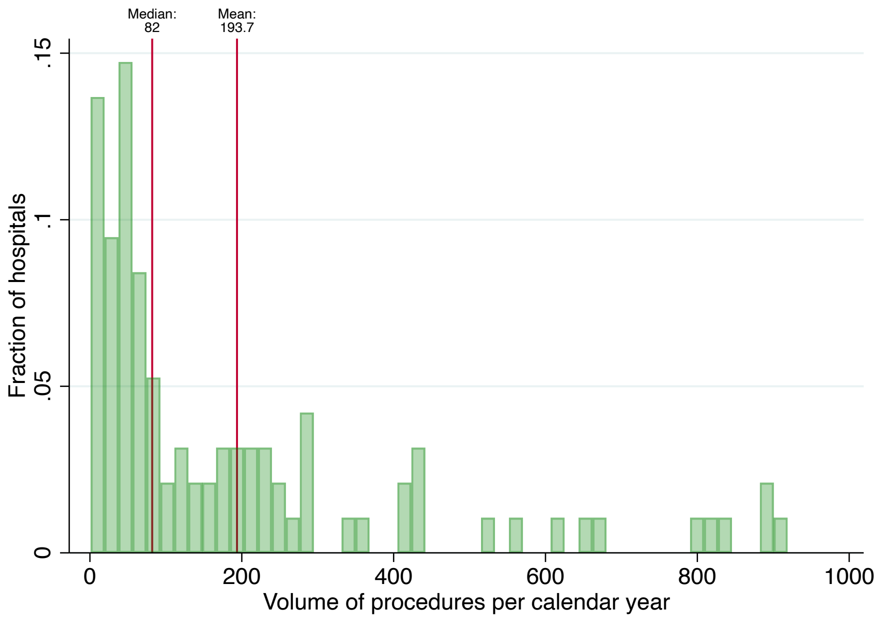

Figure 1: Empirical Distribution of Hospital Volume per Year

Figure 2 plots the relationship between hospital volume and fraction of patients dying within 30 days of procedure for a given hospital-year combination. The solid line represents fit from a fractional polynomial regression. The depiction of this

[6] Appendix A.1 provides a detailed description of all variables included in the analysis.

simple correlation between hospital volume and mortality appears to be in line with "practice makes perfect" hypothesis, although the trend seems to reverse at very high annual volumes.

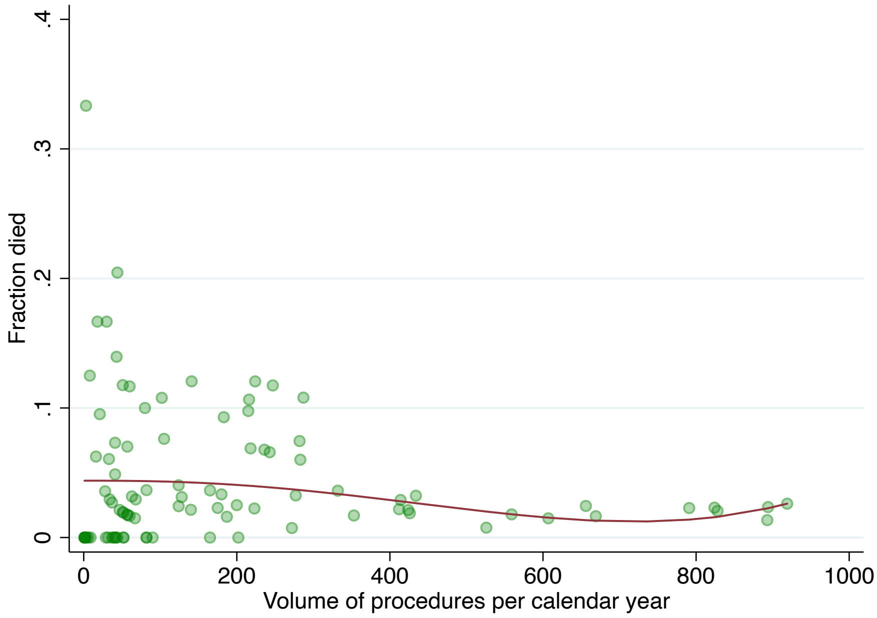

Figure 2: 30-day Mortality and Hospital Volume

Notes: Dots represent hospital/year combinations. Solid line represents fit from a fractional polynomial regression.

Figure 3 depicts the relationship between hospital volume and 30-day readmission rates for each hospital-year. As in the previous graph, the solid line represents fit from a fractional polynomial regression. A similar pattern is observed, in which hospitals performing higher volume of procedures in a given calendar year seem to have lower readmission rates until a certain volume threshold, where the trend reverses again.

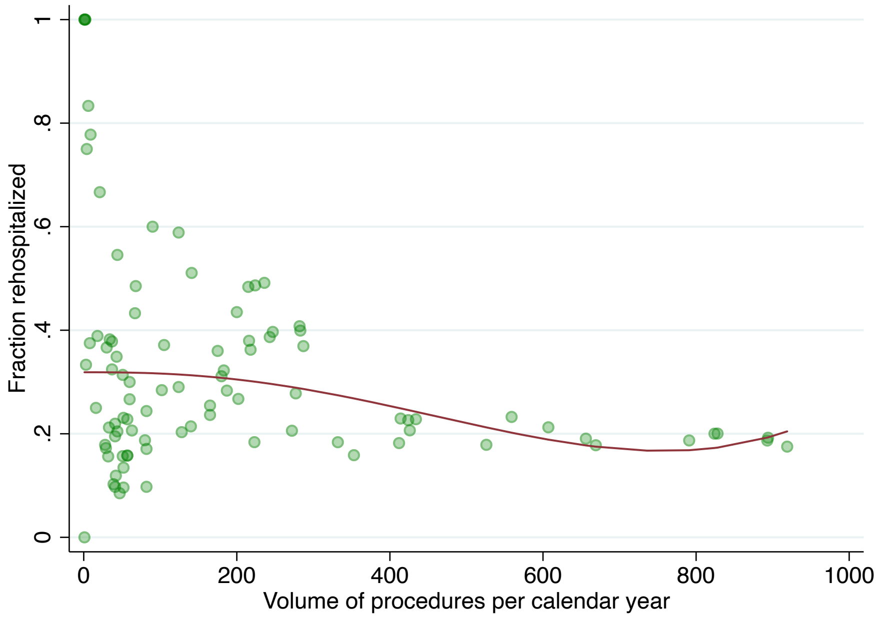

Figure 3: 30-day Readmissions and Hospital Volume

Notes: Dots represent hospital/year combinations. Solid line represents fit from a fractional polynomial regression.

A significant part of the debate within the volume-outcome literature is focused on the fact whether the observed differences rather reflect quality differences between hospitals, or if variations in volume of procedures within hospitals have effects on outcomes – as predicted by practice makes perfect hypothesis. To provide a *prima facie* summary of differences between hospitals, their outcomes and case mix, I follow Hamilton and Hamilton (1997), who first calculate the average number of procedures performed per 12-month period (each calendar year) at each hospital. The average volume is then used to split the providers into three groups: low volume hospitals performing less than 40 procedures per year (≈ 25% of sample), average volume hospitals performing between 40 and 228 procedures (≈ 50% of sample) and high volume performing more than 228 procedures per year (≈ 25% of sample).

Table 1: Betweeen Hospital Comparison of Outcomes by Volume

	Averages			
Average number of procedures performed	Readmission rate (%) (1)	Mortality rate (%) (2)	Number of comorbidities (3)	Charlson index (4)
Low (< 40)	36.9	7.3	2.60	3.13
Average (40-228)	32.1	5.7	2.76	3.37
High (> 228)	22.1	2.6	2.96	3.58
Significance test:				
p-value	0.000	0.000	0.000	0.000

Note: p-value is from a test of the null hypothesis of equality of means across categories.

Table 1 provides the overview of differences in readmission rate, mortality rate, the average number of comorbidities and Charlson index of patients undergoing the procedure. Looking at the first column, hospitals performing less than 40 procedures per year have an average readmission rate of almost 37%, compared to 32% rate of hospitals performing an average volume of procedures. High volume hospitals performing more than 228 PTA stent implants per year appear to have somewhat lower rate of readmission, with an average of 22%. A similar trend is observed when examining mortality rates between different volume groups. Low volume hospitals have mortality of around 7.3%, while mid volume hospitals have a slightly lower mortality rate of 5.7%. Again, high volume hospitals appear to perform the best, with an average mortality rate of around 2.6%. Interestingly, the differences in outcomes do not seem to be driven by differences in case mix between the volume groups. High volume hospitals with lower readmission and mortality rates also have patients with significantly more comorbidities and higher Charlson index. In summary, the descriptive properties of the sample confirm what is depicted in figures 2 and 3 and provide some indication about high volume hospitals performing better.

To provide insights on within-hospital differences, the volume of procedures in the past 12 months prior to the procedure date τ at a hospital is calculated and compared to the average number of procedures for the given hospital over the entire sample. Hospital performance is then divided into three groups – hospitals at time τ performing at 20% or less below their average number of procedures per 12 months (approximately 20% of the sample), hospitals within ±20% of their average 12-month procedure volume (≈ 60% of the sample), and hospitals performing more than 20% of volume at date τ compared to their average (≈ 20% of the sample). Based on the ranking, the differences between the mean outcomes for each hospital within 12-month period at time τ and the mean outcome for the hospital are calculated. Table 2 summarizes the comparisons by groups of hospitals as defined in Table 1.

Table 2: Within-Hospital Comparison of Outcomes by Volume Change

	Period τ difference from hospital sample average			
Percentage difference in period τ volume from hospital sample average	Readmission rate (pp) (1)	Mortality rate (pp) (2)	Number of comorbidities (3)	Charlson index (4)
Panel A. Low-volume hospitals				
Below (< −20%)	−0.19	3.34*	−0.24*	−0.37*
Average (20% to 20%)	0.38	0.03	−0.04*	−0.08*
Above (> 20%)	−3.00*	−2.64*	−0.04	−0.13*
Panel B. Average-volume hospitals				
Below (< −20%)	−5.43*	−2.57*	0.24*	0.28*
Average (20% to 20%)	0.02	0.20*	−0.02*	−0.02*
Above (> 20%)	1.60*	0.65*	−0.14*	−0.19*
Panel C. High-volume hospitals				
Below (< −20%)	−0.34*	−0.36*	−0.04*	−0.09*
Average (20% to 20%)	0.50*	0.03*	0.02*	0.02*
Above (> 20%)	0.01	0.24*	0.01*	0.03*

Notes: * significantly different from zero at the 5 percent level. pp = percentage points.

The evidence provided is somewhat surprising given the proposed direction of volume-outcome relationship. Only low-volume hospitals performing 20% or more procedures above their sample average appear to have lower readmission rate by approximately three percentage points. A similar outcome is observed for 30-day mortality rate, where low-volume hospitals during periods of 20% higher volume compared to their average have mortality rates lower by almost 2.6 percentage points. On the contrary, for average- and high-volume hospitals, an increase in procedural volume by 20% or more compared to the hospital average was associated with an increase in mortality by 0.6 and 0.2 percentage points respectively. There are also differences in case mix, as shown in columns (3) and (4). For average-volume hospitals, periods of average and above-average volume are associated with a decrease in the number of comorbidities by 0.02 and 0.14 respectively, while periods of low volume are associated with an increase of 0.24. For high-volume hospitals, below-average periods are associated with less ill patients, while average- and above-average volume periods are associated with patients with slightly more comorbidities and a higher Charlson index. The observation that both rehospitalization and mortality rate increase in average and high-volume hospitals which perform above their sample average may be explained by increased personnel demands, leading to fatigue and overwork. Patterns shown in table 2 indicate some evidence of selective referral, where less sick patients appear to be sorted into hospitals performing higher volume of procedures at the time, compared to their sample average. This could be explained by a scenario where hospitals with extra capacity for procedures are admitting fewer sick patients.

Another possible mechanism driving selective referral may be related to perceived quality of health care providers by patients. Under this scenario, hospitals with higher quality ratings may be able to attract more patients, increasing the volume of procedures performed. Figure 4 depicts a simple relationship between the index of perceived quality by patients and the volume of procedures performed by hospital in the respective calendar year. Despite the fact that there is more variation between the two, higher quality seems to be correlated with higher volume. Similarly as in previous graphs, the relationship reverses at high volumes approaching 700 and more procedures per year.

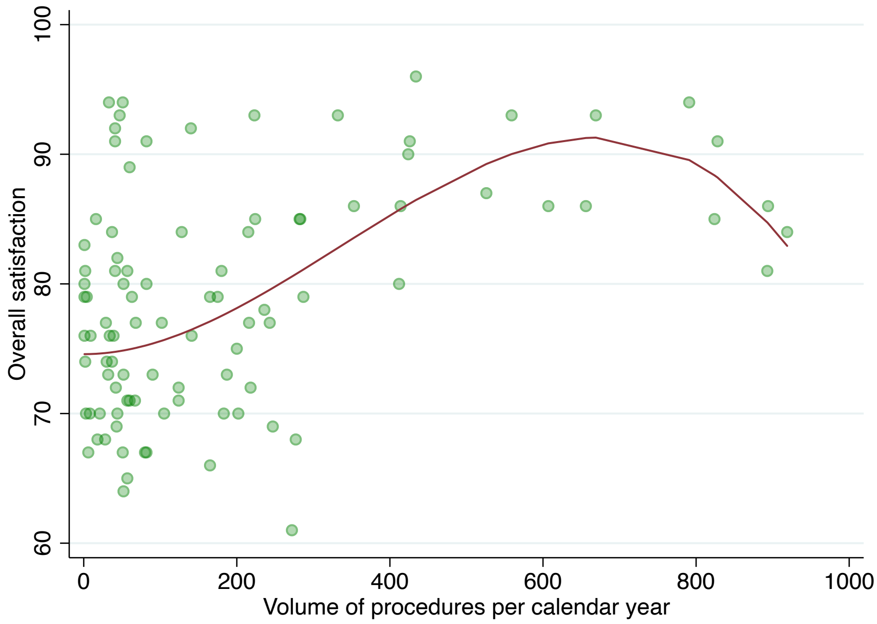

Figure 4: Perceived Quality and Hospital Volume

Notes: Dots represent hospital/year combinations. Solid line represents fit from a fractional polynomial regression.

4 Set-up of the analysis

As outlined in previous sections, analysis of causal effects between hospital volume and performance is complicated by several factors possibly affecting the relationship, such as referral patterns or differences in case mix. Other unobservables determining the outcomes may also play an important role. The analysis employed in this paper estimates the effects in a joint framework, in which hospital choice is correlated with patient outcomes using a discrete mixture of unobserved heterogeneity.

4.1 Hospital choice

The first part of the framework employs a conditional logit model, in which patients face a choice of a hospitals for the PTA procedure. The traditional conditional logit model was first introduced by McFadden (1973) and became a standard in the analysis of choice behavior. The model rests on a set of assumptions, mainly on the independence from irrelevant alternatives (IIA), which stipulates that introduction of further alternatives in the choice does not affect the odds between the initial set. The IIA assumption makes the standard conditional logit model rather restrictive in many applications. Over time, this led to adoption of a more flexible mixed logit model, for which the random coefficients are assumed to follow parametric distributions, most commonly the normal or log-normal distribution. However, as with any parametric approach, the correct specification of the underlying distribution is crucial and also carries drawbacks, as pointed out by Train (2008) and Pacifico (2013). A different approach relaxing assumptions of the standard conditional logit model is based on a non-parametric discrete mixture of unobserved heterogeneity as proposed by Heckman and Singer (1984), where the agent's preferences are allowed to vary depending on a class membership. These specifications were first explored in discrete choice models by Swait (1994) and Bhat (1997), and later employed in analysis of labor supply by Pacifico (2013).

The model for hospital choice is based on a utility function as in Varkevisser et al. (2012), who use conditional and mixed logit models to analyze the effects of quality ratings on hospital choice of patients undergoing angioplasty in the Netherlands. Under the public health insurance scheme, patients in Slovakia are free to choose a health care provider. Thus, for each patient in the dataset, there is no restriction in their choice set based on distance to hospital, and each hospital performing at least one PTA procedure during the particular year of a patient's procedure is assumed to be a possible alternative. The utility of patient i at hospital j is determined by the distance of the hospital and its perceived quality and is formally defined as:

$$U_{ij} = \delta d_{ij} + \sum_{k=1}^{n} \phi_k H_{kj} + \epsilon_{ij} \tag{1}$$

where d_{ij} denotes distance from patient's residence to hospital j, H_{kj} is a vector of observed attributes of hospital j observed by patient and ϵ_{ij} represents the error term. The main idea behind choice models with a discrete mixture of unobserved heterogeneity is that agents are sorted in a number of classes which differ in their preferences.[7] Thus, the conditional probability that patient i within a class ν chooses hospital j is defined as:

$$\Pr(J_i = j \mid z_{ij}, \gamma_\nu) = \frac{\exp(z'_{ij}\gamma_\nu)}{\sum_{k=1}^{J_i} \exp(z'_{ik}\gamma_\nu)} \tag{2}$$

where z' denotes vector of hospital attributes including travel distance, J_i is a choice set of available hospitals for a given patient i and γ_ν is a vector of preference parameters allowed to differ between classes.

4.2 Duration models

To estimate the effects of volume on 30-day mortality and readmission rates, I use a duration analysis framework, where both observed and unobserved characteristics are allowed to affect the hazard rates. Similar duration models were used previously by Hamilton and Hamilton (1997) to analyze the hospital volume-outcome relationship, or by van Ours et al. (2013) to study drug use dynamics. Figure 5 plots the empirical hazard rates for both 30-day mortality and readmissions.

[7] The number of classes is usually determined in a stepwise manner, where the researcher starts estimation with two classes and increases the number until convergence of the log-likelihood function fails to improve.

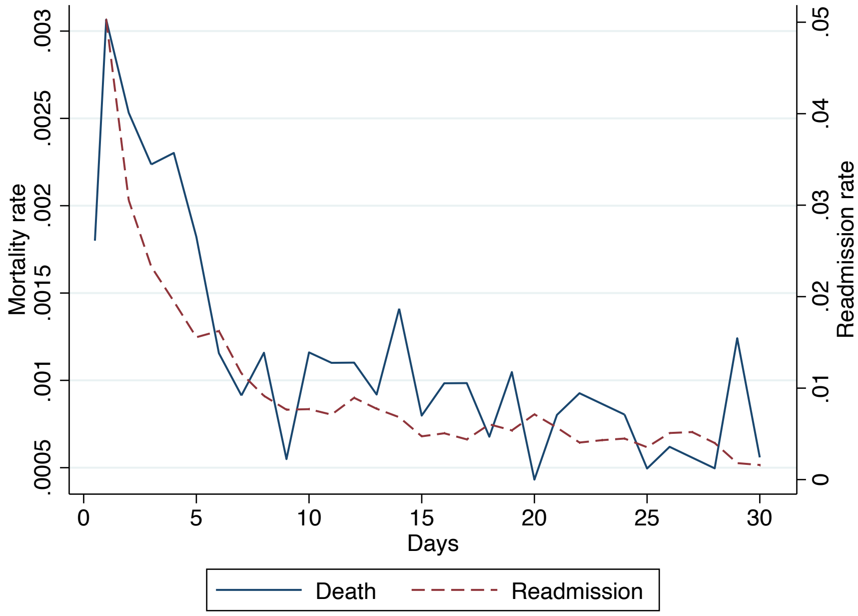

Figure 5: Empirical Hazard Rates

Both hazard rates clearly peak within first days after discharge and substantially decay for the remainder of the 30-day window following the PTA procedure, with both rates displaying some degree of non-monotonicity. Duration models allow for a variety of parametric and non-parametric specifications to describe the functional form of underlying baseline hazard. I use a mixed proportional hazard (MPH) rate model with a flexible (stepwise) duration dependence, which should approximate the baseline hazard rate with a sufficient number of intervals. More formally, the hazard rate of death within 30 days of the PTA procedure (omitting the individual subscripts) at time t_d conditional on vector of observed characteristics x[8], volume of procedures $V_{h\tau}$ and unobserved characteristics υ is specified as:

$$\theta_d(t_d \mid x, V_{h\tau}, \upsilon) = \lambda_d(t_d) \exp(x'\beta_d + V_{h\tau}\zeta_d + \alpha_{hd} + \upsilon) \tag{3}$$

where $V_{h\tau}$ is a measure of procedural volume for past 12 months at the given hospital h prior to the admission of the patient at calendar time τ, α_{hd} represents hospital-specific fixed effect and υ represents a random effect capturing the

[8]For notational simplicity, the vector x also includes an indicator for the hospital.

Heckman and Singer (1984) type of unobserved heterogeneity. $\lambda_d(t_d)$ represents individual duration dependence, which is flexibly modelled using a step function:

$$\lambda_d(t_d) = \exp\left(\sum_k \lambda_{d,k} I_k(t_d)\right) \tag{4}$$

where $k(= 1, ..., K)$ is a subscript for day-intervals and $I_k(t_d)$ are time-varying dummy variables for subsequent day-intervals when the event (death) occurs. The intervals are defined for days 0-2, 3-6, 7-10, 11-16, 17-22 and more than 22 days. I estimate a constant and normalize $\lambda_0 = 0$. The conditional density function of completed durations until death can be written as:

$$f(t_d \mid x, V_{h\tau}, \upsilon) = \theta_d(t_d \mid x, V_{h\tau}, \upsilon) \exp\left(-\int_0^{t_d} \theta_d(s \mid x, V_{h\tau}, \upsilon)\, ds\right) \tag{5}$$

The hazard rate of being rehospitalized at time t_r, conditional on observed characteristics x, volume of procedures $V_{h\tau}$ and unobserved characteristics ω is specified similarly as the hazard rate of death:

$$\theta_r(t_r \mid x, V_{h\tau}, \omega) = \lambda_r(t_r) \exp(x'\beta_r + V_{h\tau}\zeta_r + \alpha_{hr} + \omega) \tag{6}$$

where λ_r represents stepwise duration dependence with intervals defined for days as previously. The model for 30-day readmissions is specified as a competing risk, where the competing event is the risk of death. Therefore, the conditional density function of completed durations until readmission or death is defined as:

$$\begin{aligned} f(t_r \mid x, V_{h\tau}, \omega, \upsilon) = & [\theta_r(t_r \mid x, V_{h\tau}, \omega) + \theta_d(t_r \mid x, V_{h\tau}, \upsilon)] \\ & \exp\left(-\int_0^{t_r} [\theta_r(u \mid x, V_{h\tau}, \omega) + \theta_d(u \mid x, V_{h\tau}, \upsilon)]\ du\right) \end{aligned} \tag{7}$$

Durations where neither readmission nor death is observed within 30 days are considered right-censored.

4.3 Specification of unobserved heterogeneity

Selection of hospital and patient outcomes may be jointly affected by unobserved characteristics. For example, the unobserved health status of a patient may affect choice of a hospital as well as the probability of readmission or death after the procedure. It is conceivable that patients in a worse health status are referred to hospitals with a higher perceived quality, or to specialized cardiac centers in order to obtain better care. These patients may be characterized by willingness to travel further to a hospital with better quality ratings. Also, a prior belief and expectations about receiving higher quality care, as well as interactions during hospital stay may have an effect on post-surgical outcomes.[9] If such unobserved factors are not accounted for in the empirical framework, estimates capturing the relationship between volume of procedures and hospital performance are likely to be biased. Specification of heterogeneity distribution where the unobserved factors in both hospital choice and patient outcomes are allowed to be correlated between each other should capture these hidden differences.

I assume that the random effects υ, ω and latent classes γ_ν come from a discrete mixing distribution G, where each of the components has two points of support. The joint density function for 30-day mortality is then specified as:

$$g(J = j, t_d \mid z_j, x, V_{h\tau}) = \int_{\gamma_\nu} \int_{\upsilon} \Pr(J = j \mid z_j, \gamma_\nu) f(t_d \mid x, V_{h\tau}, \upsilon)\, dG(\gamma_\nu, \upsilon) \tag{8}$$

while for 30-day readmissions it is defined as:

$$\begin{aligned} g(J = j, t_r \mid z_j, x, V_{h\tau}) = &\int_{\gamma_\nu} \int_{\omega} \int_{\upsilon} \Pr(J = j \mid z_j, \gamma_\nu) \\ &f(t_r \mid x, V_{h\tau}, \omega, \upsilon)\, dG(\gamma_\nu, \omega, \upsilon) \end{aligned} \tag{9}$$

Therefore, the full mixing distribution yields four and eight possible combinations respectively, each describing two types of patients with different hazard rates of death and readmission (high hazard rate and low hazard rate) and two types of patients with different preferences for hospital choice. The probabilities associated with four mass points of the unobserved heterogeneity distribution for hospital

[9] Associations between patient satisfaction and outcomes were explored for example by Boulding et al. (2011)

choice and 30-day mortality are denoted as:

$$\begin{aligned} &\Pr(\upsilon = \upsilon_1, \gamma_\nu = \gamma_{\nu_1}) = p_1, \Pr(\upsilon = \upsilon_1, \gamma_\nu = \gamma_{\nu_2}) = p_2 \\ &\Pr(\upsilon = \upsilon_2, \gamma_\nu = \gamma_{\nu_1}) = p_3, \Pr(\upsilon = \upsilon_2, \gamma_\nu = \gamma_{\nu_2}) = p_4 \end{aligned} \tag{10}$$

with $p_4 = 1-p_1-p_2-p_3$, where p_n $(n = 1, ..., 4)$ is assumed to follow a multinomial logistic distribution:

$$p_n = \frac{\exp(\alpha_n)}{\sum_n \exp(\alpha_n)},\ n = 1, ..., 4 \tag{11}$$

with α_4 normalized to zero. The probabilities associated with eight mass points of the distribution for hospital choice and 30-day readmissions are defined as:

$$\begin{aligned} &p_1 = \Pr(\upsilon = \upsilon_1, \omega = \omega_1, \gamma_\nu = \gamma_{\nu_1}), p_2 = \Pr(\upsilon = \upsilon_1, \omega = \omega_1, \gamma_\nu = \gamma_{\nu_2}) \\ &p_3 = \Pr(\upsilon = \upsilon_1, \omega = \omega_2, \gamma_\nu = \gamma_{\nu_1}), p_4 = \Pr(\upsilon = \upsilon_1, \omega = \omega_2, \gamma_\nu = \gamma_{\nu_2}) \\ &p_5 = \Pr(\upsilon = \upsilon_2, \omega = \omega_1, \gamma_\nu = \gamma_{\nu_1}), p_6 = \Pr(\upsilon = \upsilon_2, \omega = \omega_1, \gamma_\nu = \gamma_{\nu_2}) \\ &p_7 = \Pr(\upsilon = \upsilon_2, \omega = \omega_2, \gamma_\nu = \gamma_{\nu_1}), p_8 = \Pr(\upsilon = \upsilon_2, \omega = \omega_2, \gamma_\nu = \gamma_{\nu_2}) \end{aligned} \tag{12}$$

and modelled as a multinomial logit similarly as in equation 11, with p_8 defined analogously to p_4.[10] The likelihood functions are optimized over all unknown parameters using a method of maximum likelihood.

5 Results

The following pages present results of the empirical model outlined in the previous section of the paper. The results of the model are also presented as independent processes, i.e. without the correlated unobserved heterogeneity and then estimated jointly. The main parameter of interest is the variable capturing the volume of procedures at the hospital within last 12 months prior to the current patient, defined as natural logarithm $\ln(V_{h\tau})$ to follow previous publications. Other covariates included in the hospital choice equation are distance to the hospital, overall readmission rate, total number of patients per year, overall satisfaction and dummy

[10]Note that equations 10 and 12 describe full mixing distributions, where only some of the probability mass points may be identified during estimation.

variables for a university hospital and specialized cardiac centre. For the 30-day mortality and readmissions, region of residence[11] age and gender of the patient, type of procedure (depending on the affected artery), categorical Charlson index and also respective dummy variables for the 16 comorbidity categories of Charlson index are included as explanatory variables. Furthermore, the outcome equations also include fixed effects for the hospital.[12]

5.1 30-day mortality

Table 3 summarizes the results for 30-day mortality after the PTA procedure. Column (1) of Panel A. reports estimates of the conditional logit model for hospital choice without latent classes for comparison. The results show that patients are more likely to choose health care providers with higher ratings, as well as hospitals whcih admit more patients per year. They are also more likely to choose specialized cardiology centers and hospitals closer to their residence.

Columns (2) and (3) in Panel B. report estimates from the duration model with and without hospital fixed effects. Not surprisingly, the results reveal that older patients are more likely to die within 30 days. There are also differences between arteries involved in PTA procedures. Compared to the reference category of PTA involving peripheral arteries of extremities, procedures involving celiac, mesenteric or carotid artery carry a significantly higher risk of death. Panel C. reports parameters of the duration dependence, revealing that the risk of death is highest within first two days after the procedure. Turning to the main parameter of interest, the coefficient capturing the effects of hospital volume on 30-day mortality is negative and statistically significant in the model without fixed effects, suggesting a positive effect of higher volume of procedures on 30-day mortality after the PTA procedure. In numerical terms, the point estimate reveals that an increase of procedural volume by 10% is associated with approximately 2.4% decrease of mortality rate.[13] Once hospital fixed effects are included, the estimated coefficient nearly halves in magnitude to 1.4%, but remains statistically significant

[11] Alternative specifications of the model with variables capturing median income of patient's residence were also estimated (results not reported), finding the same conclusions.

[12] See appendix A.1 for a full description of variables.

[13] $(\exp(-0.26 \times \ln(1.1)) - 1) \times 100 \approx -2.4$

at the 10% level. This is similar to results by Hamilton and Hamilton (1997) in their analysis of volume-outcome relationship of patients undergoing hip surgery. The fact that the volume-outcome effects disappear once time-invariant hospital characteristics are controlled for suggests that the results to a large extent reflect differences in quality between low and high volume hospitals. The penultimate row of table 3 reports log-likelihood values of respective models. A likelihood-ratio test comparing models (2) and (3) suggests that model with fixed effects provides a significant improvement.

Finally, column (4) reports estimates from the model where selection of hospital and mortality outcomes are estimated jointly. Panel A. summarizes estimates related to hospital choice. There are two classes of patients who differ in their preferences for hospital choice. The results are best interpreted by comparison of the two classes between each other and within the context of how health care is organized in most countries. The first group of patients is more likely to travel farther to the hospital of their choice, and also is more likely to choose university affiliated hospital or a specialized cardiac center. This likely reflects the fact that university hospitals and cardiac centers are located in large cities and therefore require travel for patients from rural areas. This group of individuals is also more concerned about the quality of hospital, and also has a higher mortality rate. For the other group of patients, the distance to hospital is more of a problem, while they also show a lower preference for university hospitals and specialized cardiac centers. Coefficients related to unobserved heterogeneity distribution also reveal that this group is represented by approximately 36% of patients and also has a zero mortality rate. The most striking difference between independent models and the joint model is that once the correlated unobservables between hospital choice and mortality are taken into account, the effect of hospital volume on outcomes further decreases to 0.9% and is no longer statistically significant at the conventional significance levels. This result suggests that ignoring selectivity in hospital choice overestimates the volume-outcome relationship. A likelihood-ratio test comparing log-likelihood values between models with and without unobserved heterogeneity confirms that the model with correlated error structure provides a better fit.

Table 3: Parameter Estimates 30-day Mortality Rate

	(1)		(2)		(3)		(4)			
Panel A. Hospital choice										
Rehospitalization rate	0.02	(5.4)*					−0.08	(9.1)***	0.28	(16.0)***
Overall satisfaction	0.03	(11.5)***					0.08	(18.5)***	−0.02	(2.3)**
Number of patients per year	0.00	(1.7)*					−0.01	(3.2)**	0.02	(5.7)***
University hospital	−0.01	(0.3)					0.82	(8.8)***	−1.64	(11.4)***
Cardiac center	1.34	(43.3)***					2.12	(42.8)***	−0.70	(4.1)***
Distance to hospital	−0.02	(106.8)***					−0.02	(61.4)***	−0.09	(16.8)***
Panel B. Mortality										
$ln(V_{h\tau})$			−0.26	(6.9)***	−0.15	(1.7)*	−0.09	(0.9)		
Other characteristics										
Age			0.05	(10.3)***	0.05	(9.9)***	0.05	(9.8)***		
Female			0.09	(1.0)	0.05	(0.6)	0.08	(0.9)		
PTA of the arch of aorta			0.49	(1.1)	0.43	(0.9)	0.44	(0.9)		
PTA of carotid and vertebral arteries			1.38	(10.3)***	1.09	(7.6)***	1.03	(6.9)***		
PTA of renal arteries			−0.98	(2.1)**	−1.14	(2.5)**	−0.96	(2.0)**		
PTA of celiac or mesenteric artery			1.42	(5.2)***	1.31	(4.7)***	1.48	(4.9)***		
Duration dependence										
Days 3-6			−0.66	(5.1)***	−0.66	(5.0)***	−0.59	(4.5)***		
Days 7-10			−1.32	(8.3)***	−1.31	(8.2)***	−1.18	(7.4)***		
Days 11-16			−1.19	(8.8)***	−1.18	(8.7)***	−1.01	(7.3)***		
Days 16-22			−1.46	(9.9)***	−1.45	(9.8)***	−1.24	(8.3)***		
Days 22+			−1.58	(11.6)***	−1.56	(11.5)***	−1.31	(9.5)***		
Unobserved heterogeneity										
υ_1			−7.84	(18.5)***	−6.08	(7.0)***	−4.39	(4.3)***		
$\upsilon_1 - \upsilon_2$							−∞			
α_1							0.55	(12.4)***		
Hospital FE	-		No		Yes		Yes			
Unobserved heterogeneity	No		No		No		Yes			
p-value hospital FE[a]	-		-		0.000		0.000			
p-value comorbidities[b]	-		0.000		0.000		0.000			
-Log likelihood	25,402.1		3,984.2		3,950.0***		27,789.4***			
Observations	16,599		16,599		16,599		16,599			

Notes: Absolute t statistics in parentheses. FE = fixed effects. Coefficients for regional and comorbidity indicator variables not reported. Full set of results for the main specification is available in Appendix B.1, Table B1.

[a] Test of the null hypothesis that the coefficients on the hospital indicators are jointly zero.

[b] Test of the null hypothesis that the coefficients on the comorbidity indicators are jointly zero.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

5.2 30-day readmissions

Results for 30-day readmissions are presented in the same fashion, where column (1) of table 4 repeats estimates from the model of hospital choice without latent classes, while columns (2) and (3) present estimates from the competing risk model with and without fixed effects. Finally, column (4) reports estimates from the joint correlated model.

A similar story emerges as reported in the previous section once the effects of hospital volume are estimated with and without fixed effects. Patients are at the highest risk of being rehospitalized within the first two days after discharge, after which the hazard rate decreases, as shown in Panel C. The model without fixed effects finds a highly significant effect of hospital volume on 30-day readmissions, suggesting that an increase in volume of PTA procedures is related to lower rates of readmission. However, once fixed differences between hospitals are accounted for, the coefficient shrinks from -0.18 to -0.08, but remains statistically significant. The penultimate row of table 4 confirms that the model with fixed effects presents a significant improvement over the model in column (2). Looking at column (4), where the correlated error structure is taken into account, the coefficient further decreases to -0.05 and again loses statistical significance. The distribution of unobserved heterogeneity reveals that there are two types of patients – the first type represented by 80% of patients is characterized by a higher willingness to travel farther to a hospital and a higher preference for an university affiliated hospital or a specialized cardiology center. This group is also characterized by a higher readmission rate and a lower mortality rate. The other group represented by 20% of patients shows less preference for university hospitals and special centres, is less concerned about quality of hospitals in terms of readmission rates, and also has a lower willingness to travel farther. Furthermore, these patients have zero rates of readmission compared to the first group and higher mortality rates. The fact that the coefficient related to hospital volume further decreases and loses statistical significance once the correlation between unobserved components in hospital choice and hazard rates of readmission is accounted for suggests evidence of selective referral.

Table 4: Parameter Estimates 30-day Readmission Rate

	(1)		(2)		(3)		(4)			
Panel A. Hospital choice										
Rehospitalization rate	0.02	(5.4)***					0.01	(0.9)	0.16	(3.6)***
Overall satisfaction	0.03	(11.5)***					0.05	(16.1)***	0.05	(2.7)**
Number of patients per year	0.00	(1.7)*					0.00	(1.3)	0.06	(4.6)***
University hospital	−0.01	(0.3)					0.33	(6.2)***	−6.85	(6.1)***
Cardiac center	1.34	(43.3)***					1.67	(45.9)***	−3.10	(5.1)***
Distance to hospital	−0.02	(106.8)***					−0.02	(83.7)***	−0.37	(7.0)***
Panel B. Readmissions										
$ln(V_{h\tau})$			−0.18	(12.5)***	−0.08	(2.0)**	−0.05	(1.2)		
Other characteristics										
Age			0.01	(5.8)***	0.01	(5.0)***	0.01	(5.0)***		
Female			0.06	(1.9)*	0.05	(1.5)	0.03	(1.0)		
PTA of the arch of aorta			0.16	(1.3)	0.27	(2.2)**	0.24	(1.9)*		
PTA of carotid and vertebral arteries			0.54	(10.7)***	0.55	(10.4)***	0.43	(7.7)***		
PTA of renal arteries			−0.72	(5.9)***	−0.65	(5.3)***	−0.56	(4.3)***		
PTA of celiac or mesenteric artery			0.19	(1.4)	0.16	(1.2)	0.11	(0.7)		
Duration dependence										
Days 3-6			−0.06	(1.3)	−0.04	(0.9)	0.06	(1.4)		
Days 7-10			−0.71	(13.4)***	−0.68	(12.8)***	−0.51	(9.5)***		
Days 11-16			−1.14	(20.8)***	−1.10	(20.1)***	−0.89	(16.0)***		
Days 16-22			−1.37	(23.0)***	−1.33	(22.3)***	−1.08	(17.7)***		
Days 22+			−1.80	(29.5)***	−1.76	(28.8)***	−1.47	(23.5)***		
Unobserved heterogeneity										
ω_1			−3.31	(23.1)***	−3.04	(5.1)***	−2.35	(3.2)**		
υ_1							−6.27	(4.4)***		
$\omega_1 - \omega_2$							$-\infty$			
$\upsilon_1 - \upsilon_2$							−4.49	(3.2)***		
α_1							1.40	(41.8)***		
Hospital FE	-		No		Yes		Yes			
Unobserved heterogeneity	No		No		No		Yes			
p-value hospital FE[a]	-		-		0.000		0.000			
p-value comorbidities[b]	-		0.000		0.000		0.000			
-Log likelihood	25,402.1		23,692.3		23,504.6***		47,736.6***			
Observations	16,599		16,599		16,599		16,599			

Notes: Absolute t statistics in parentheses. FE = fixed effects. Competing risk of death estimates not reported. Coefficients for regional and comorbidity indicator variables not reported. Full set of results for the main specification is available in Appendix B.1, Table B2.

[a] Test of the null hypothesis that the coefficients on the hospital indicators are jointly zero.

[b] Test of the null hypothesis that the coefficients on the comorbidity indicators are jointly zero.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

It is conceivable that the first type of unobserved heterogeneity individuals with higher rates of readmission and lower mortality rates, and also higher preference for university hospitals and specialized cardiac centres represents patients more concerned about their health. These are likely willing to seek the highest quality health care provided by specialized cardiac centers and university hospitals. Comparison of log-likelihood values between models with and without unobserved heterogeneity confirms that the joint model provides a better fit.

5.3 Sensitivity analysis

Estimates presented in the previous section considered all-cause readmission within 30-day period as an outcome. A potential drawback of this measure is that not all readmissions are necessarily related to complications after PTA with stent implant. The most common risks associated with the procedure are narrowing of the affected artery, forming of blood clots within stents and bleeding. Other complications include acute myocardial infarction, arterial damage, kidney problems, stroke or abnormal heart rhythms. Therefore, only readmissions with discharge diagnoses related to skin or wound lesions, stroke, certain kidney syndromes, circulatory system and complications following surgical or medical care are considered in the sensitivity analysis. Full list of diagnoses included is provided in Appendix A.2, table A1. Table 5 summarizes the results of the model. For the sake of brevity, only the main parameters of interest are reported.

As in previous sections, column (1) presents estimates of the competing risk model without hospital fixed effects, column (2) with hospital fixed effects and column (3) reports the joint correlated model with unobserved heterogeneity. The results to a great extent resemble all-cause 30-day readmissions, despite being based on a more restrictive set of discharge diagnoses. Without hospital fixed effects, the parameter estimate of hospital volume is negative and highly significant, with a point estimate of -0.14. Once the differences between hospitals are accounted for, the estimate shrinks to -0.10, but remains significant. In the correlated model, the estimated effect further decreases to -0.06 and is no longer significant at the conventional levels.

Table 5: Parameter Estimates 30-day Readmission Rate – Sensitivity Analysis

	(1)		(2)		(3)	
Hospital volume						
$\ln(V_{h\tau})$	-0.14	$(7.4)^{***}$	-0.10	$(2.2)^{**}$	-0.06	(1.3)
Unobserved heterogeneity						
ω_1	-4.36	$(24.0)^{***}$	-3.54	$(4.8)^{***}$	-3.86	$(8.1)^{***}$
$\omega_1 - \omega_2$					$-\infty$	
υ_1					-9.43	$(7.6)^{***}$
$\upsilon_1 - \upsilon_2$					-8.01	$(6.5)^{***}$
α_1					-1.15	$(34.4)^{***}$
Hospital FE	No		Yes		Yes	
Unobserved heterogeneity	No		No		Yes	
p-value hospital FE[a]	-		0.000		0.000	
p-value comorbidities[b]	0.000		0.000		0.000	
-Log likelihood	17,815.1		17,636.5***		41,714.4***	
Observations	16,599		16,599		16,599	

Notes: Absolute t statistics in parentheses. FE = fixed effects. Coefficients for hospital choice, personal characteristics, regional, comorbidity, and hospital fixed effects indicator variables not reported. Competing risk of death estimates not reported.
[a] Test of the null hypothesis that the coefficients on the hospital indicators are jointly zero.
[b] Test of the null hypothesis that the coefficients on the comorbidity indicators are jointly zero.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

6 Conclusions

Relationship between volume of hospital procedures and outcomes is a relatively well researched topic in the health economics literature, with a large number of studies finding a positive relationship between the two. However, only few acknowledge the difficulty in attributing causal effects to the proposed "practice makes perfect" hypothesis. This paper analyzes the universe of patients undergoing percutaneous transluminal angioplasty with stent implant in Slovakia over 6 year period. The main contributions of this paper are twofold. First, the empirical model used exploits the variation in the volume of procedures over time and controls for potential selectivity in hospital choice by using a correlated discrete factor approach. Compared to previous studies, selection into hospital is estimated jointly with patient outcomes measured as 30-day readmission and mortality. Second, the analysis is focused on short-term post-procedural outcomes rather than on outcomes related to in-hospital stay.

To some extent, conclusions from the empirical models confirm those previously reported by Hamilton and Hamilton (1997). Once the effect of volume on outcomes is estimated solely on between-hospital variation of procedures, the relationship appears highly significant and supportive of "practice makes perfect" hypothesis. However, if controls for time-invariant hospital characteristics are included, the effect reduces in size. This suggests that the volume-outcome relationship to some extent reflects differences in quality between health care providers. Still, even after controlling for the hospital fixed effects, there is a statistically significant, positive effect of higher volume on outcomes. However, the estimated positive volume-outcome relationship further shrinks and is no longer statistically significant once selection of patients into hospital is taken into account in the joint correlated model. The results also reveal that the risk of death or readmission is highest within the first days after the procedure, confirming that focusing solely on in-hospital outcomes may yield biased results. The findings are also robust to the specification when 30-day readmissions are considered on a more restrictive set of diagnoses.

Most of the results in the hospital volume-outcome literature appear to be specific to the type of procedure or specialization, where positive effects of larger volume on outcomes were found, for example, for advanced cancer surgeries, while no effects were reported for cardiac revascularizations. Within the context of cardiac procedures, the findings presented in this paper are in contrast to those reported by Phillips et al. (1995), who find that hospitals performing more PTA procedures had significantly better outcomes, measured as lower mortality and shorter LOS. However, compared to the results presented in this study, their analysis relied on cross-sectional data covering one year and focused on in-hospital outcomes. Furthermore, the analysis did not control for hospital fixed effects and potential selectivity.

All in all, these findings should serve as a caution for strict proponents of EBHR limits and regionalization of hospital care. If the volume-outcome relationship indeed mostly reflects selective referral and differences in quality between hospitals for certain procedures, a closure and transformation of low-volume providers into high-volume providers might not lead to better health outcomes for patients, but to worse access to health care.

References

Avdic, D., P. Lundborg, and J. Vikström (2019). Estimating returns to hospital volume: Evidence from advanced cancer surgery. *Journal of Health Economics 63*, 81–99.

Bannay, A., C. Chaignot, P.-O. Blotière, M. Basson, A. Weill, P. Ricordeau, and F. Alla (2016). The best use of the Charlson comorbidity index with electronic health care database to predict mortality. *Medical Care 54*(2), 188–194.

Barker, D., G. Rosenthal, and P. Cram (2011). Simultaneous relationships between procedure volume and mortality: do they bias studies of mortality at specialty hospitals? *Health Economics 20*(5), 505–518.

Bhat, C. R. (1997). An endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science 31*(1), 34–48.

Boulding, W., S. W. Glickman, M. P. Manary, K. A. Schulman, and R. Staelin (2011, January). Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. *The American Journal of Managed Care 17*(1), 41–48.

Farley, D. E. and R. J. Ozminkowski (1992). Volume-outcome relationships and in-hospital mortality: The effect of changes in volume over time. *Medical Care 30*(1), 77–94.

Gaynor, M., H. Seider, and W. B. Vogt (2005). The volume-outcome effect, scale economies, and learning-by-doing. *American Economic Review 95*(2), 243–247.

Hamilton, B. H. and V. H. Hamilton (1997). Estimating surgical volume-outcome relationships applying survival models: accounting for frailty and hospital fixed effects. *Health Economics 6*(4), 383–395.

Hannan, E. L., J. Kilburn, H, J. F. O'Donnell, H. R. Bernard, E. P. Shields, M. L. Lindsey, and A. Yazici (1992). A longitudinal analysis of the relationship between in-hospital mortality in New York State and the volume of abdominal aortic aneurysm surgeries performed. *Health Services Research 27*(4), 517–542.

Heckman, J. and B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*(2), 271–320.

Luft, H. (1980). The relation between surgical volume and mortality: an exploration of causal factors and alternative models. *Medical Care 18*(9), 940–959.

Luft, H. S., J. P. Bunker, and A. C. Enthoven (1979). Should operations be regionalized? *New England Journal of Medicine 301*(25), 1364–1369.

Luft, H. S., S. S. Hunt, and S. C. Maerki (1987). The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health Services Research 22*(2), 157–182.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York, NY, USA: Academic Press New York.

Pacifico, D. (2013). On the role of unobserved preference heterogeneity in discrete choice models of labour supply. *Empirical Economics 45*(2), 929–963.

Phillips, K. A., H. S. Luft, and J. L. Ritchie (1995). The association of hospital volumes of percutaneous transluminal coronary angioplasty with adverse outcomes, length of stay, and charges in California. *Medical Care 33*(5), 502–514.

Quan, H., V. Sundararajan, P. Halfon, A. Fong, B. Burnand, J.-C. Luthi, L. D. Saunders, C. A. Beck, T. E. Feasby, and W. A. Ghali (2005). Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical Care 43*(11), 1130–1139.

Swait, J. (1994). A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. *Journal of Retailing and Consumer Services 1*(2), 77–89.

Train, K. E. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling 1*(1), 40–69.

Tsai, A. C., M. Votruba, J. F. P. Bridges, and R. D. Cebul (2006). Overcoming bias in estimating the volume-outcome relationship. *Health Services Research 41*(1), 252–264.

van den Berg, G., B. van der Klaauw, and J. van Ours (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics 22*(1), 211–241.

van der Klaauw, B. and J. C. van Ours (2013). Carrot and stick: how re-employment bonuses and benefit sanctions affect exit rates from welfare. *Journal of Applied Econometrics 28*(2), 275–296.

van Ours, J. C. and J. Williams (2012). The effects of cannabis use on physical and mental health. *Journal of Health Economics 31*(4), 564–577.

van Ours, J. C., J. Williams, D. Fergusson, and L. J. Horwood (2013). Cannabis use and suicidal ideation. *Journal of Health Economics 32*(3), 524–537.

Varkevisser, M., S. A. van der Geest, and F. T. Schut (2012). Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands. *Journal of Health Economics 31*(2), 371–378.

Appendix A

A.1 Definition of variables

- **Hospital choice**
  - Readmission rate: ratio of patients discharged from a given hospital and rehospitalized within 30 days for a diagnosis in the same group, to all patients hospitalized in the same hospital for the respective diagnosis group.
  - Overall satisfaction: A composite index of 12 standardized indicators related to perceived quality of health care by patient. Includes subjective ratings of health care, patient-doctor communication, accommodation, cleanliness, nutrition and subjective assessment of treatment. Collected by insurance companies using telephone surveys, anonymous on-line questionnaires or by standard post service. Hospitals and health care providers are sorted according to the rating, with the best hospital receiving 100 points, while the lowest ranked hospital receiving 0. Remaining hospitals receive points relative to the best ranked provider.
  - Patients per year: total number of patients hospitalized at the hospital during a calendar year.
  - University hospital: hospital is an university affiliated (teaching) hospital.
  - Cardiac center: Hospital is a specialized cardiac center.
  - Distance to hospital: calculated as a point-to-point distance in kilometres from patient's residence to the hospital.
- **Outcomes**
  - Duration until death: measured in days since the date of procedure. Patients dying on the same day are coded with a duration of 0.5.
  - Duration until readmission: measured in days since discharge from hospital. Readmissions considered if they occur at least 24 hours or later after discharge.
- **Other characteristics**
  - Age: age in years.

- Region: indicator variables for region of residence.
- Residence: observed on the level of town/city, in the case of larger cities on city district.
- Female: indicator variable for female gender.
- Type of procedure: indicator variables for type of PTA procedure depending on the involved artery. Baseline category - PTA of peripheral arteries of extremities and pelvic arteries.
- Charlson index: Charlson comorbidity index, categorical variable. Baseline category: 0-1.
- Comorbidity indicator variables: 16 indicator variables for Charlson comorbidity index categories. Metastatic cancer and HIV/AIDS categories are merged into one due to low number of outcomes observed for HIV/AIDS category.
- $\ln(V_{h\tau})$: Natural logarithm of volume of procedures performed in the preceding 12 months at a given hospital, prior to the current patient's procedure.

A.2 Readmission diagnoses

Table A1: List of Diagnoses Considered in Sensitivity Analysis

Diagnosis code	Description
A40, A41	Streptococcal sepsis, other sepsis
G45	Transient cerebral ischemic attacks and related syndromes
J95	Intraoperative and postprocedural complications and disorders of respiratory system, not elsewhere classified
L02	Cutaneous abscess, furuncle and carbuncle
L03	Acute lymphangitis
L08	Other local infections of skin and subcutaneous tissue
L76	Intraoperative and postprocedural complications of skin and subcutaneous tissue
L98	Other disorders of skin and subcutaneous tissue, not elsewhere classified
N01, N04	Rapidly progressive nephritic syndrome, nephritic syndrome
N17, N19	Acute kidney failure, unspecified kidney failure
R58	Hemorrhage, not elsewhere classified
R60	Edema, not elsewhere classified
I10-I16	Hypertensive diseases
I20-I25	Ischemic heart diseases
I26-I28	Pulmonary heart disease and diseases of pulmonary circulation
I30-I52	Other forms of heart disease
I60-I69	Cerebrovascular diseases
I70-I79	Diseases of arteries, arterioles and capillaries
I80-I89	Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified
I95-I99	Other and unspecified disorders of the circulatory system
R00-R09	Symptoms and signs involving the circulatory and respiratory systems
T80-T88	Complications of surgical and medical care, not elsewhere classified
T98	Sequelae of other and unspecified effects of external causes
Z48	Encounter for other postprocedural aftercare

Appendix B

B.1 Additional results and tables

Table B1: Parameter Estimates 30-day Mortality Rate

	(1)		(2)		(3)		(4)			
Panel A. Hospital choice										
Rehospitalization rate	0.02	(5.4)*					−0.08	(9.1)***	0.28	(16.0)***
Overall satisfaction	0.03	(11.5)***					0.08	(18.5)***	−0.02	(2.3)**
Number of patients per year	0.00	(1.7)*					−0.01	(3.2)**	0.02	(5.7)***
University hospital	−0.01	(0.3)					0.82	(8.8)***	−1.64	(11.4)***
Cardiac center	1.34	(43.3)***					2.12	(42.8)***	−0.70	(4.1)***
Distance to hospital	−0.02	(106.8)***					−0.02	(61.4)***	−0.09	(16.8)***
Panel B. Mortality										
$ln(V_h)$			−0.26	(6.9)***	−0.15	(1.7)*		−0.09	(0.9)	
Other characteristics										
Trnavský region			0.52	(2.9)**	0.04	(0.2)		0.42	(2.1)**	
Trenčiansky region			0.27	(1.4)	0.14	(0.7)		−0.56	(2.7)**	
Nitriansky region			0.26	(1.3)	0.18	(0.9)		0.06	(0.3)	
Žilinský region			0.51	(2.8)**	0.44	(1.7)*		0.53	(2.3)**	
Banskobystrický region			0.40	(2.2)**	0.45	(1.7)*		0.51	(2.2)**	
Prešovský region			0.19	(0.9)	0.24	(0.7)		−0.29	(0.9)	
Košický region			0.44	(2.2)**	0.61	(1.7)*		0.25	(0.8)	
Age			0.05	(10.3)***	0.05	(9.9)***		0.05	(9.8)***	
Female			0.09	(1.0)	0.05	(0.6)		0.08	(0.9)	
PTA of the arch of aorta			0.49	(1.1)	0.43	(0.9)		0.44	(0.9)	
PTA of carotid and vertebral arteries			1.38	(10.3)***	1.09	(7.6)***		1.03	(6.9)***	
PTA of renal arteries			−0.98	(2.1)**	−1.14	(2.5)**		−0.96	(2.0)**	

Continued on next page

Table B1 – Continued From Previous Page

	(1)	(2)		(3)		(4)	
PTA of celiac or mesenteric artery		1.42	(5.2)***	1.31	(4.7)***	1.48	(4.9)***
Comorbidities							
Recent AMI		0.09	(0.7)	0.13	(1.0)	0.18	(1.3)
Congestive heart failure		0.45	(4.4)***	0.45	(4.5)***	0.45	(4.2)***
Peripheral vascular disease		−0.69	(6.1)***	−0.62	(5.4)***	−0.55	(4.6)***
Cerebrovascular disease		−0.89	(9.3)***	−0.75	(7.6)***	−0.69	(6.7)***
Dementia		0.66	(4.4)***	0.62	(4.1)***	0.80	(4.8)***
COPD		−0.11	(1.0)	−0.11	(1.0)	−0.08	(0.7)
Rheumatoid disease		0.32	(1.6)	0.34	(1.7)*	0.49	(2.4)**
Peptic ulcer disease		−0.27	(1.4)	−0.27	(1.4)	−0.42	(2.1)**
Mild liver disease		−0.01	(0.1)	0.01	(0.1)	−0.04	(0.3)
Diabetes		0.10	(1.0)	0.12	(1.1)	0.15	(1.3)
Diabetes + complications		0.16	(1.4)	0.16	(1.4)	0.12	(0.9)
Hemiplegia or paraplegia		−0.26	(0.6)	−0.19	(0.5)	−0.09	(0.2)
Renal disease		0.29	(2.2)**	0.26	(1.9)*	0.29	(2.0)**
Cancer		0.09	(0.7)	0.09	(0.8)	0.02	(0.1)
Severe renal disease		−0.09	(0.1)	−0.08	(0.1)	0.08	(0.1)
Metastatic cancer/AIDS		0.62	(2.1)**	0.57	(2.0)*	0.97	(2.9)**
Duration dependence							
Days 0-2		−0.66	(5.1)***	−0.66	(5.0)***	−0.59	(4.5)***
Days 3-6		−1.32	(8.3)***	−1.31	(8.2)***	−1.18	(7.4)***
Days 11-16		−1.19	(8.8)***	−1.18	(8.7)***	−1.01	(7.3)***
Days 16-22		−1.46	(9.9)***	−1.45	(9.8)***	−1.24	(8.3)***
Days 22+		−1.58	(11.6)***	−1.56	(11.5)***	−1.31	(9.5)***
Unobserved heterogeneity							
υ_1		−7.84	(18.5)***	−6.08	(7.0)***	−4.39	(4.3)***
$\upsilon_1 - \upsilon_2$						$-\infty$	

Continued on next page

Table B1 – Continued From Previous Page

	(1)	(2)	(3)	(4)	
α_1				0.55	(12.4)***
Hospital FE	-	No	Yes	Yes	
Unobserved heterogeneity	No	No	No	Yes	
p-value hospital FE[a]	-	-	0.000	0.000	
p-value comorbidities[b]	-	0.000	0.000	0.000	
-Log likelihood	25,402.1	3,984.2	3,950.0***	27,789.4***	
Observations	16,599	16,599	16,599	16,599	

Notes: Absolute t statistics in parentheses.

[a] Test of the null hypothesis that the coefficients on the hospital indicators are jointly zero.

[b] Test of the null hypothesis that the coefficients on the comorbidity indicators are jointly zero.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Table B2: Parameter Estimates 30-day Readmission Rate

	(1)		(2)		(3)		(4)			
Panel A. Hospital choice										
Rehospitalization rate	0.02	(5.4)***					0.01	(0.9)	0.16	(3.6)***
Overall satisfaction	0.03	(11.5)***					0.05	(16.1)***	0.05	(2.7)**
Number of patients per year	0.00	(1.7)*					0.00	(1.3)	0.06	(4.6)***
University hospital	−0.01	(0.3)					0.33	(6.2)***	−6.85	(6.1)***
Cardiac center	1.34	(43.3)***					1.67	(45.9)***	−3.10	(5.1)***
Distance to hospital	−0.02	(106.8)***					−0.02	(83.7)***	−0.37	(7.0)***
Panel B. Readmissions										
$ln(V_h)$			−0.18	(12.5)***	−0.08	(2.0)**		−0.05	(1.2)	
Other characteristics										
Trnavský region			0.27	(3.9)***	−0.04	(0.5)		0.05	(0.7)	

Continued on next page

Table B2 – Continued From Previous Page

	(1)		(2)		(3)		(4)	
Trenčiansky region			0.57	(8.8)***	0.18	(2.3)**	−0.10	(1.3)
Nitriansky region			0.05	(0.7)	0.08	(1.1)	−0.23	(2.9)**
Žilinský region			0.63	(9.7)***	0.13	(1.4)	0.10	(1.2)
Banskobystrický region			0.33	(5.0)***	−0.00	(0.1)	0.00	(0.0)
Prešovský region			0.38	(5.7)***	0.12	(1.0)	−0.07	(0.6)
Košický region			0.37	(5.6)***	0.25	(2.0)**	0.13	(1.1)
Age			0.01	(5.8)***	0.01	(5.0)***	0.01	(5.0)***
Female			0.06	(1.9)*	0.05	(1.5)	0.03	(1.0)
$ln(V_h)$			−0.18	(12.5)***	−0.08	(2.0)**	−0.05	(1.2)
PTA of the arch of aorta			0.16	(1.3)	0.27	(2.2)**	0.24	(1.9)*
PTA of carotid and vertebral arteries			0.54	(10.7)***	0.55	(10.4)***	0.43	(7.7)***
PTA of renal arteries			−0.72	(5.9)***	−0.65	(5.3)***	−0.56	(4.3)***
PTA of celiac or mesenteric artery			0.19	(1.4)	0.16	(1.2)	0.11	(0.7)
Comorbidities								
Recent AMI			0.00	(0.0)	0.02	(0.4)	0.03	(0.5)
Congestive heart failure			0.18	(4.5)***	0.19	(4.7)***	0.15	(3.6)***
Peripheral vascular disease			−0.60	(13.1)***	−0.54	(11.8)***	−0.51	(10.4)***
Cerebrovascular disease			−0.55	(14.8)***	−0.49	(13.0)***	−0.42	(10.8)***
Dementia			0.31	(4.2)***	0.27	(3.6)***	0.33	(4.1)***
COPD			0.01	(0.3)	0.03	(0.7)	0.04	(1.0)
Rheumatoid disease			−0.18	(2.1)**	−0.16	(1.9)*	−0.11	(1.1)
Peptic ulcer disease			0.11	(1.8)*	0.11	(1.7)*	0.08	(1.2)
Mild liver disease			−0.01	(0.3)	0.00	(0.1)	−0.03	(0.6)
Diabetes			−0.03	(0.8)	−0.02	(0.6)	0.00	(0.1)
Diabetes + complications			0.26	(6.1)***	0.26	(6.0)***	0.22	(4.8)***
Hemiplegia or paraplegia			−0.06	(0.4)	−0.08	(0.6)	0.01	(0.0)
Renal disease			0.29	(6.0)***	0.27	(5.6)***	0.28	(5.5)***

Continued on next page

Table B2 – Continued From Previous Page

	(1)	(2)		(3)		(4)	
Cancer		0.07	(1.6)	0.08	(1.9)*	0.11	(2.3)**
Severe renal disease		0.05	(0.2)	0.05	(0.2)	0.18	(0.8)
Metastatic cancer/AIDS		0.34	(2.6)**	0.35	(2.6)**	0.36	(2.6)**
Duration dependence							
Days 0-2		−0.06	(1.3)	−0.04	(0.9)	0.06	(1.4)
Days 3-6		−0.71	(13.4)***	−0.68	(12.8)***	−0.51	(9.5)***
Days 11-16		−1.14	(20.8)***	−1.10	(20.1)***	−0.89	(16.0)***
Days 16-22		−1.37	(23.0)***	−1.33	(22.3)***	−1.08	(17.7)***
Days 22+		−1.80	(29.5)***	−1.76	(28.8)***	−1.47	(23.5)***
Unobserved heterogeneity							
ω_1		−3.31	(23.1)***	−3.04	(5.1)***	−2.35	(3.2)**
υ_1						−6.27	(4.4)***
$\omega_1 - \omega_2$						$-\infty$	
$\upsilon_1 - \upsilon_2$						−4.49	(3.2)***
α_1						1.40	(41.8)***
Hospital FE	-	No		Yes		Yes	
Unobserved heterogeneity	No	No		No		Yes	
p-value hospital FE[a]	-	-		0.000		0.000	
p-value comorbidities[b]	-	0.000		0.000		0.000	
-Log likelihood	25,402.1	23,692.3		23,504.6***		47,736.6***	
Observations	16,599	16,599		16,599		16,599	

Notes: Absolute t statistics in parentheses. [a] Test of the null hypothesis that the coefficients on the hospital indicators are jointly zero.
[b] Test of the null hypothesis that the coefficients on the comorbidity indicators are jointly zero.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Table B3: Descriptive Statistics

Variable	Mean	SD	Min	Max
Hospital choice				
Rehospitalization rate	8.32	3.54	2.71	21.89
Overall satisfaction	79.61	8.44	61	96.00
Number of patients per year	24.14	17.56	0.53	86.45
University hospital	0.19	0.39	0	1
Cardiac center	0.19	0.39	0	1
Distance to hospital	182.94	115.27	0.60	511.60
Hospital volume				
$ln(V_{h\tau})$	5.82	0.97	-4.61	6.83
Other characteristics				
Rehospitalized	0.25	0.43	0	1
Died	0.03	0.18	0	1
Days until rehospitalization	9.30	10.22	1	30
Days until death	11.61	4.39	1	30
Trnavský region	0.12	0.32	0	1
Trenčiansky region	0.11	0.32	0	1
Nitriansky region	0.10	0.30	0	1
Žilinský region	0.10	0.30	0	1
Banskobystrický region	0.13	0.33	0	1
Prešovský region	0.15	0.36	0	1
Košický region	0.16	0.36	0	1
Age	66.69	11.45	0	100
Female	0.35	0.48	0	1
PTA of peripheral arteries of extremities	0.63	0.48	0	1
PTA of the arch of aorta	0.02	0.14	0	1
PTA of carotid and vertebral arteries	0.31	0.46	0	1
PTA of renal arteries	0.03	0.18	0	1
PTA of celiac or mesenteric artery	0.01	0.11	0	1
Comorbidities				
Recent AMI	0.12	0.32	0	1
Congestive heart failure	0.19	0.39	0	1
Peripheral vascular disease	0.77	0.42	0	1
Cerebrovascular disease	0.47	0.50	0	1

Continued on next page

Table B3 – Continued From Previous Page

Variable	Mean	SD	Min	Max
Dementia	0.04	0.19	0	1
COPD	0.19	0.39	0	1
Rheumatoid disease	0.04	0.19	0	1
Peptic ulcer disease	0.06	0.24	0	1
Mild liver disease	0.12	0.33	0	1
Diabetes	0.35	0.48	0	1
Diabetes + complications	0.30	0.46	0	1
Hemiplegia or paraplegia	0.01	0.11	0	1
Renal disease	0.10	0.30	0	1
Cancer	0.13	0.33	0	1
Severe renal disease	0.01	0.07	0	1
Metastatic cancer/AIDS	0.01	0.10	0	1
Observations		16,599		