

Van Gestel, Raf; Müller, Tobias; Bosmans, Johan

Working Paper

Learning from failure in healthcare: Dynamic panel evidence of a physician shock effect

Discussion Papers, No. 18-09

Provided in Cooperation with:

Department of Economics, University of Bern

Suggested Citation: Van Gestel, Raf; Müller, Tobias; Bosmans, Johan (2018) : Learning from failure in healthcare: Dynamic panel evidence of a physician shock effect, Discussion Papers, No. 18-09, University of Bern, Department of Economics, Bern

This Version is available at:

<https://hdl.handle.net/10419/204909>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



^b
**UNIVERSITÄT
BERN**

Faculty of Business, Economics
and Social Sciences

Department of Economics

**Learning from failure in healthcare: Dynamic panel
evidence of a physician shock effect**

Raf Van Gestel, Tobias Müller,
Johan Bosmans

18-09

April 2018

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
<http://www.vwi.unibe.ch>

Learning from failure in healthcare: dynamic panel evidence of a physician shock effect

Raf Van Gestel^{*}, Tobias Müller[†], Johan Bosmans[‡]

ABSTRACT

Procedural failures of physicians or teams in interventional healthcare may positively or negatively predict subsequent patient outcomes. We identify this effect by applying (non-)linear dynamic panel methods to data from the Belgian Transcatheter Aorta Valve Implantation (TAVI) registry containing information on the first 860 TAVI procedures in Belgium. We find that a previous death of a patient positively and significantly predicts subsequent survival of the succeeding patient. We find that these learning from failure effects are not long-lived and that learning from failure is transmitted across adverse events.

Keywords: Physician behavior, Learning, Failure

JEL: I10, I13, I18, C93

^{*} Corresponding author. Department of Applied Economics, Erasmus University of Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands and Department of Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium. Mail: vangelstel@ese.eur.nl, Tel.: +31 010-408 12 72

[†] University of Bern, Department of Economics, Schanzeneckstrasse 1, 3001 Bern, Switzerland. Mail: tobias.mueller@vwi.unibe.ch

[‡] Department of Cardiology, University of Antwerp, Wilrijkstraat 10, 2560 Edegem, Belgium.

1. INTRODUCTION

Physician learning is an umbrella term covering multiple types of learning, forgetting and knowledge transfer. The literature on learning for physicians typically focuses on specific case-studies and the identification of different types of learning. Most notably, a distinction is made between physician experience, economies of scale and human capital depreciation (Hockenberry and Helmchen, 2014; Van Gestel *et al.*, 2016). In Van Gestel *et al.* (2016), these three types of learning have been further investigated with a focus on patient subgroups. In this paper, we build on this work by looking at performance responses to a failure. An adverse event may, either positively or negatively, disrupt health provider performance. A failure may have an impact on performance through three primary channels. Firstly, a failure might alarm and shock a provider. As a result, in a subsequent procedure, the provider might be better prepared and/or better motivated to obtain a positive outcome. This response might arise because of, among others, loss aversion. Second, a failure may provide the physician with more information on specific aspects of the procedure. Lastly, subsequent to a failure, patients with different characteristics may be selected to undergo the procedure. Although clearly a selection effect, this might indicate that physicians learn to more appropriately select patients over time. Hence, the three channels may all represent certain aspects of provider learning from failure. Throughout the literature, endogeneity problems hamper inference of learning effects with reverse causality and risk selection over time as main validity threats (e.g. Gaynor, 2005; Hentschker and Mennicken, 2016). In this paper, our empirical strategy directly aims at addressing potential endogeneity issues when estimating learning from failure effects.

The role of the physician is, evidently, widely acknowledged to be important in the production of health. Physician incentives and behavior, but also characteristics, are expected to be predictive of patient health outcomes. Aspects like education and specialization (Jollis *et al.*, 1996), adherence to guidelines (Ward, 2006) and experience (Hockenberry and Helmchen, 2014) have all been shown to affect patients' health outcomes. At a more aggregated level, physician supply and contract-types (contracted vs. municipality GP's in Aakvik *et al.* (2006)) have been shown to be correlated with mortality rates (Or *et al.*, 2005; Sundmacher *et al.*, 2011; Iizuka, 2016).

As a result, a number of initiatives target provider performance. Firstly, these initiatives may raise physicians' information and skills through continued medical education (Cervero *et al.*, 2015), traineeships, etc. (Haynes *et al.*, 1995). Secondly, market-based interventions change incentives for patients and physicians which also influences subsequent performance. For example, publicly available report cards (Kolstad, 2013) and pay for performance programs (Li, 2014) target incentives through both intrinsic motivation and market-based stimuli. In this paper, we contribute to this literature by exploring if previous failures (e.g. adverse events like patient mortality and stroke) affect subsequent physician performance. This research may inform policy makers on the scope to introduce future information or awareness campaigns. In addition, this analysis can also serve as a starting point to further explore underlying reasons for such shock effects.

Learning from failure is important at all levels of the healthcare sector. Nation-wide failed health reforms should inform future reforms and organizations are expected to foster failure driven learning even though they might face barriers to do so (Oberlander, 2007; Edmondson, 2004). At the provider level, learning from failure might be encouraged by the information provided by the failure itself (e.g. physicians may learn more from risky patients), by the incentive to avoid malpractice claims (Pänthofer, 2017) or more generally by loss aversion. This loss aversion may

relate to income, patient break-ups because of low performance and/or adverse events (Rizzo, 2003; Hareli, 2007). However, failure may not only stimulate learning, it could also result in more failures. This result, typically found in the organizational literature, occurs when for example failed (business) projects negatively influence a team (Shepherd, 2013). Psychologically, personal goal failure may lead to negative affective states and may therefore translate into negative subsequent outcomes (Jones, 2013). Also, physician inertia may contribute to continued failure. A failure to respond to adverse outcomes may stem from habit formation and from the reluctance to adjust treatment practice because of sizeable search and learning costs (Janakiraman, 2008). Whereas the idea of learning from failure for physicians is related to research on physician inertia in pharmaceutical prescriptions, we apply the idea of previous experience and inertia to a more interventional setting.

The identification of the learning from failure effect on physician performance relies on previous experiences as patient mortality for a physician depends on mortality of the previous patient(s). Estimation of such learning from failure effects imposes two main econometric challenges: First, in the context of binary response fixed effects (FE) dynamic panel data models, the well-known incidental parameter problem and Nickell bias lead to possibly inconsistent coefficient estimates (for an overview see Lancaster, 2000). Second, with irregular spacing of panel data, FE may not appropriately be accounted for in dynamic models (McKenzie, 2001). In our data, this irregular spacing occurs because the time span between Transcatheter Aorta Valve Implantation (henceforth TAVI) procedures differs within hospitals. That is, we only observe provider activity when a patient is actually treated in a given facility and therefore the time span between procedures is irregularly spaced between patients.

As a remedy to the incidental parameter problem and the Nickell bias, we apply the bias-corrected FE estimator proposed by De Vos *et al.* (2015) which has shown to have superior small sample properties compared to classical GMM-type estimators and FE linear probability models. In addition to the bias-correction approach, we estimate the learning from failure effects using the split-panel jackknife FE estimator recently proposed by Dhaene and Jochmans (2015). We also discuss the irregular spacing of our data in which the dependency between observations can be modelled as an AR(1)-process.

In this paper, we provide evidence for substantial learning from failure effects. We find that if the previous patient died, the probability to die within one month is associated to decrease by 5 to 11%-points for the next patient. Although different non-linear dynamic panel methods provide slightly different results, all specifications provide qualitatively similar results. We also find minor evidence for a transmission mechanism of shocks between adverse events. A previous stroke, a common complication during or shortly after the TAVI procedure, is correlated with a lower likelihood of dying.

In the remainder of this paper, we discuss the background of our application to TAVI and our data in section 2. In section three we focus on the methods to measure the effect of learning from failure. In section 4, we present our results after which we discuss and conclude in section 5.

2. THE INTRODUCTION OF TRANSCATHETER AORTA VALVE IMPLANTATION (TAVI)

The application in this paper considers the introduction and evolution of the TAVI procedure in Belgium for which the first procedure in a Belgian hospital has taken place in 2007. With a TAVI procedure, the defunct aortic valve is replaced with an artificial replacement valve through means

of a minimally invasive catheterization. In our sample period the procedure was not reimbursed and was therefore only financed by physicians and hospitals for patients that were anatomically inoperable and that were unable to bear the cost themselves. Later, in 2015, TAVI became reimbursed under certain conditions.

We use register data for patients from 2007, including the first patient undergoing TAVI, to the beginning of 2012 in 20 hospitals. As such, our data is able to describe the performance of physicians for the full introduction period of TAVI. In each hospital only one team performs this TAVI procedure and during the sample period total workload for TAVI was limited to about one day a week. Furthermore, team composition hardly varies over the sample period 2007-2012. Information is available on a wide range of patient specific characteristics (see Table I for details) and we have access to hospital identifiers. However, the data does not contain information about additional hospital (e.g. hospital budget) or surgeon-team characteristics (e.g. years of experience). Note that with the exception of “Ejection fraction”, all variables in our data set are binary indicators. Approval was obtained from the institutional ethics committees for data collection at the participating hospitals. On average, 1-month mortality amounts to about 9% with proportions ranging between 5 to 27% between hospitals. In addition, Table II gives an overview on the total number of patients treated over the different years and Table I shows the total number of patients treated at each of the different hospitals over the entire sample period.

--- Insert Table I here ---

--- Insert Table II here ---

Table III provides first descriptive evidence for the learning from failure effect for interventional care. Descriptive numbers are provided for four common adverse events: mortality, having a stroke,

renal failure and pacemaker implantation. Cerebrovascular stroke is a common complication during the procedure because of, e.g. blood or calcium cloths that are dislodged during the procedure (for a technical explanation see Ghelam *et al.*, 2016). Also, renal failure and pacemaker implantation regularly take place during or shortly after the TAVI procedure (Eggebrecht *et al.*, 2011). The estimates in the “Lag=0” columns contain the probabilities for the adverse event if the previous patient did not suffer from the corresponding adverse event. The column “Lag=1” shows the likelihood of the adverse event after an adverse event. For 1-month mortality and stroke, the probability of suffering from a stroke or dying within one-month is substantially lower if the previous patient exhibited the same adverse event pointing towards learning from failure effects. For renal failure and pacemaker, point estimates have the opposite sign and are statistically insignificant. Throughout the next sections, we will thoroughly scrutinize these preliminary descriptive findings. Since only mortality and stroke exhibit interesting patterns and because only for mortality there are subsequent failures, we will focus our attention entirely on mortality.¹

--- Insert Table III here ---

3. METHODS

3.1. Empirical specification and learning curves

The literature on learning curves in health distinguishes between three types of learning (see e.g. Hockenberry and Helmchen (2014); Van Gestel *et al.* (2016)): cumulative experience (CE), economies of scale (EOS) and human capital depreciation (HCD). Learning from cumulative

¹ The fact that there are no consecutive failures, which is apparent from the Lag=1 column in Table III makes it impossible to estimate non-linear models (because coefficients tend to infinity as they represent multiplicative effects).

experience refers to the idea that treating an additional patient generally improves physician (or team) performance. When referring to economies of scale, we capture the fact that higher volume providers usually have better infrastructure (e.g. equipment, staff) and more standardized procedures. Lastly, the human capital depreciation hypothesis states that provider performance decreases with longer temporal distance to previous procedures. In its most simple form, this leads to the specification in equation (1) which can empirically be estimated using standard regression techniques:

$$Outcome_{i,h,t} = \beta_0 + \beta_1 HCD_{i,h} + \beta_2 CE_{i,h} + \beta_3 EOS_{h,t} + \alpha'_1 X_{i,h,t} + \theta'_1 H_h + \varepsilon_{i,h,t} \quad (1)$$

In the empirical strategy, our data are considered a pseudo-panel dataset in which the individual patients are consecutive observations on provider and hospital performance. In equation (1), i , h and t denote individual i in hospital h in year t ². The outcome variables in our analysis are binary mortality indicators. Furthermore, we control for a vector of background characteristics and comorbidities contained in $X_{i,h,t}$ and hospital fixed effects H_h . Lastly, $\varepsilon_{i,h,t}$ is a typical error term which in case of LPM's is non-normally distributed. Consequently, we use heteroscedasticity-robust standard errors in all our model specifications.

$$Outcome_{i,h,t} = \beta_0 + \beta_1 Outcome_{i-1,h,t} + \beta_2 HCD_{i,h} + \beta_3 CE_{i,h} + \beta_4 EOS_{h,t} + \alpha'_1 X_{i,h,t} + \theta'_1 H_h + \varepsilon_{i,h,t} \quad (2)$$

In equation (2), we extend (1) by adding a lag of the outcome variable as additional independent variable to the model. Note here that the outcome, say mortality, of patient i in hospital h in period t is regressed on the mortality indicator of the previous patient $i-1$ which was treated just before

² Note that this notation changes in the robustness section on irregular spacing where the time units are specific to each individual observation and the differences are expressed in days.

patient i in the same hospital h in period t . β_1 is therefore capturing the learning from failure effect. Furthermore, we include the learning variables from equation (1) to overcome the potential omitted variable bias resulting from a likely correlation between the lagged outcome and the learning indicators. In fact, without the inclusion of the typical learning effects, β_1 would likely be an upper bound on the true effect³.

Because of the dynamics in equation (2), there are several issues that may obstruct estimation and inference of the failure effect. First, the binary outcomes may require the estimation of a non-linear dynamic panel model with fixed effects. In this setting the fixed effects may induce a substantial incidental parameter problem and a Nickell bias. Second, by estimating the dependence between observations, the irregular spacing in the model may cause the within-estimator to be biased because the panel unit fixed effects are not fully accounted for. These estimation issues are further discussed throughout the following sections.

3.2. Dynamic panels and the failure effect

Since the dynamic effects for the adverse events are of primary interest in this paper, our main goal is to obtain consistent and efficient estimates of the potential failure effect. Specifically, we analyze whether the patient outcome of a team or physician procedure is correlated with previous (negative) experiences. In this section we describe potential improvements over the basic FE LPM shown in equation (2).

³ I.e. it would probably overstate the positive effect of a previous failure on survival as e.g. cumulative experience is negatively correlated to both contemporaneous and lagged patient mortality pointing toward an upward bias in $\widehat{\beta}_4$.

3.2.1. Incidental parameters

In the context of binary dependent variables and fixed effects, the incidental parameters problem frequently presents a hurdle to obtain consistent and efficient estimates. Because of the incidental parameters problem, the fixed effects estimates are inconsistently estimated and this carries over into inconsistency for all other coefficients (Baltagi, 2008; Wooldridge, 2010). The inclusion of incidental parameters is even more problematic with lagged dependent variables because of the Nickell bias⁴ (Moon *et al.*, 2015). To address these endogeneity issues, difference and system GMM-type estimators are commonly used in the applied literature. For example, Salge *et al.* (2016) apply dynamic instrumental variable panel methods in relationship with quality of care. They find that infections decrease with better overall cleaning, training on infection control, hand hygiene and a favorable error-reporting environment. We refrain from applying GMM-type estimators as they regularly suffer from poor small sample properties due to weak instrument problems (see Pua, 2015; De Vos *et al.*, 2012; Bruno, 2005). Instead, to address these econometric challenges, we run a series of alternative estimation techniques. As a starting point, we estimate the learning from failure effects using simple fixed effects Linear Probability Models (henceforth FE LPM's) and fixed effects logit (henceforth FE logit) specifications. Although widely used in the applied literature, Chernozhukov *et al.* (2013; p. 546) demonstrate that the FE LPM provides inconsistent estimates of the average marginal effects in dynamic panel settings. Furthermore, applying non-linear fixed effects models (e.g. FE logit/probit) has shown to produce persistent (upward) bias in the estimated slope coefficients due to the incidental parameters problems. In fact, based on Monte

⁴ As shown in Nickell (1981), the inclusion of lagged dependent variables leads to a direct correlation of the lagged outcome with the error term rendering the FE estimator inconsistent.

Carlo simulations, Greene (2004) shows that the bias in the estimated FE logit and probit coefficients increases, the smaller the time horizon T in the data.

To address these econometric challenges, we apply in a first step the bias-corrected FE estimator proposed by De Vos *et al.* (2015) based on Everaert and Pozzi (2007)⁵ which has been shown to have superior small sample properties than the classical GMM-type estimators as it removes most of the bias in the classical FE estimator. The underlying idea behind the bias-corrected FE estimator is to estimate the bias in the (uncorrected) FE estimate by using different bootstrap samples of the original data and iteratively improve the initial estimate until reaching an unbiased estimator.

In a second step, we apply the split-panel jackknife for non-linear fixed effect models recently suggested by Dhaene and Jochmans (2015) which addresses the incidental parameters problem, while at the same time accounting for the dynamics. The method is specifically designed for moderately large panels. We apply the jackknife-based approach because we do not include the typical time fixed effects⁶ as in Fernández-Val and Weidner (2016) where the analytical bias correction is preferred. Whereas most solutions to the abovementioned difficulties are computationally complex and require analytical solutions, the jackknife is relatively straightforward to implement and performs similarly or even better than other approaches. The drawback however is the difficulty to include time trends in non-linear fixed effects models which makes it hard/impossible to correct for the overall learning curve. With different dynamics in the subpanels, the estimates of all coefficients may differ. Consequently, we provide a range of estimates.

⁵ Because the approach is designed for continuous dependent variables, we simulate the bias in AME's in a Monte Carlo Simulation.

⁶ This is simply because our constructed "panel" does not have a strict time dimension which makes it difficult to interpret and because we include a range of time trends (HCD, CE, EOS) as independent variables of interest.

The intuition underlying the split-panel jackknife is to divide each panel in smaller subpanels of consequent observations.⁷ Because the bias depends on the length T of the panel, using different panel lengths by generating subpanels and comparing the coefficient estimates of the subpanels with the estimate for the complete panel provides an estimate of the bias. Subtracting the estimated bias from the complete panel estimate generates the split-panel jackknife estimator (Dhaene and Jochmans, 2015, p. 998). One simple choice to determine the subpanel lengths also suggested in Dhaene and Jochmans (2015) is the half-panel jackknife where the panels are simply divided in two. In a dynamic setting (p. 1007) the jackknife may perform sub-optimally when the dynamics are very different in the half-panels. Subpanel estimates can be compared to test for sensitivity to differential dynamics.

3.2.2. Irregularly spaced panel data

In our setting, patients are considered as consecutive observations on provider performance. As such, because periods between patients may substantially vary (across hospitals and patients), our panel can be considered as irregularly spaced. With irregularly spaced panel data, the fixed effects approach may not fully account for the fixed effects and as a consequence, the within estimator may be biased (McKenzie, 2001; Tamm *et al.*, 2007). One approach, that will be discussed in more detail in the robustness section is to model the data as an AR(1)-process to better illustrate the dependence between patients. Although intuitively attractive, technical and practical considerations hinder the applicability of this method in our setting.

⁷ We provide a brief and intuitive summary of the split-panel jackknife. For more detailed technical information, please see Dhaene and Jochmans (2015). For more information on the technical implementation in Stata, consult the help files on `probitfe`. <https://ideas.repec.org/c/boc/bocode/s458278.html> or Cruz-Gonzalez, Fernandez-Val and Weidner (2017)

4. RESULTS: EVIDENCE ON LEARNING FROM FAILURE

4.1. Baseline estimates

As a starting point, we estimate simple fixed effects linear probability (FE LPM) and fixed effects logit (FE logit) specifications before presenting the bias-corrected FE and the split-panel jackknife estimates. Table IV below presents the FE LPM and FE logit average discrete probability effects (ADPE) of our learning from failure effects for the outcomes of one- and 24-month mortality after the TAVI procedure. Both specifications include patient- and procedure specific characteristics (see descriptive statistics in Table I), hospital fixed effects and the different learning effects described above. Remark that, especially for 24-month mortality, the lagged outcome may not yet be available to the physician at the start of the next procedure. In this case, we make the implicit assumption that there is a latent variable indicating treatment success according to the physician that closely matches the one or 24-month mortality.

Overall, we find highly significant negative coefficients on our lagged outcome variables for both mortality outcomes. In fact, our FE LPM estimates indicate that the predicted likelihood of dying in the first month after the TAVI procedure is associated to decrease by about 7.8%-points if the last patients past away. Likewise, the probability of dying 2-years after the procedure is predicted to decrease by about 7.4%-points pointing again towards strong learning from failure effects. The ADPE estimates in the FE logit specifications are of similar magnitude and thus lead to the same conclusions. These preliminary findings suggest that there is substantial room for self-correction and personal improvement after failure.

Furthermore, in line with the findings in Van Gestel *et al.* (2016), our estimates provide evidence for a significant positive learning from cumulative experience effect as treating an additional

patient is associated with a decrease in 2-year mortality of about 0.2%-points. Although seemingly a small effect, this quickly becomes large for sizeable patient samples.

--- Insert Table IV here ---

4.2. Bias-corrected FE LPM and split-panel jackknife FE probit estimates

As discussed in the methodology section, the dynamics in the model specifications above cause the errors to be correlated with the lagged dependent variables inducing a Nickell Bias on all coefficient estimates shown in Table IV. The preferable strategy therefore is to consider the dynamics and simultaneously address the incidental parameter problem. To this end, we apply the bias-corrected FE estimator suggested by De Vos *et al.* (2015) and the split-panel jackknife for probit models proposed by Dhaene and Jochmans (2015).

Table V below shows the bias-corrected FE estimates of the learning from failure effects for both mortality outcomes. The specification is based on 250 bootstrap samples and we use the burn-in initialization scheme to set the initial values of the lagged dependent variables⁸. Moreover, we allow for general heteroscedasticity of the error term using the wild bootstrap suggested by Liu (1988) and Mammen (1993) in the algorithm.

Overall, the coefficients in Table V below are qualitatively and quantitatively similar to the results in Table IV: the likelihood that the current patient dies within one- or 24-months after the TAVI procedure is associated to decrease if the last patient died. Although not significant anymore for one-month mortality, they have the same sign and order of magnitude compared to previous results. In addition, we again find evidence for significant learning from cumulative experience effects for

⁸ The burn-in initialization has the advantage (over analytic schemes) that one does not have to make specific distributional assumptions on the initial conditions in the bootstrap algorithm.

2-year mortality and human capital depreciation. Note that in contrast to the FE LPM and logit estimates above, the bias-corrected FE estimates tend to be smaller in absolute value for the mortality indicators thus pointing toward upward bias in the LPM estimates above⁹.

--- Insert Table V here ---

In a next step, we estimate the learning from failure effects using the split-panel jackknife FE probit estimator recently proposed by Dhaene and Jochmans (2015). Table VI below shows the estimated learning from failure effects for 1-month and 2-year mortality while controlling for the usual patient- and procedure-specific characteristics and also including hospital fixed effects. Overall, we again find significant evidence for the presence of a learning from failure effect. Specifically, our split-panel jackknife estimate suggest that the likelihood of survival is associated to increase by about 11%-points (resp. 7%-points) if the last patient passed away within one-month (24-months) after the TAVI procedure.

In conclusion, although point estimates between methods vary, our different estimation approaches all hint that physician (or team) performance tends to positively respond to previous failures. We observe both a significant decrease in the likelihood of short- and long-term mortality for the next patient.

--- Insert Table VI here ---

⁹ To assess the relative performance of the bias-corrected FE estimator against the FE LPM, we ran Monte Carlo simulations showing that the bias-corrected FE estimator clearly outperforms the FE LPM in terms of bias . The results may however be biased with serially correlated errors. (see Appendix A for details).

4.3. Nature and interpretation of the failure effect

The significant lagged effect may be interpreted in several ways. Firstly, we might expect that more can be learned from a failure than a success. If this would hold true, we would expect that this translates in a (slope-) shift of the “typical” learning curve in figure 1 below. In the figure, two fitted lines are shown. The black dashed line represents average mortality of patients at different levels of experience, conditional on non-mortality of the previous patient. The grey full line conditions on mortality of the previous patient. A slope-shift in this figure would mean that more is learned from early failures as opposed to later failures. I.e. after the first patient died, the mortality may significantly drop (from the black dashed to the grey full line) while this drop could be lower at higher patient numbers. However, the figure suggests that the same downward effect of a previous failure on mortality is found across all patient numbers. This points out that the effect of an adverse event is constant at all levels of experience and we could therefore tentatively conclude that instead of influencing the learning curve, an adverse event has a constant impact. This finding could point towards a sort of concentration effect.

--- Insert Figure 1 here ---

Secondly, we are interested in the persistence of the shock effect, i.e., whether longer lags on the dependent variable are still significant. However, in none of our analyses, the second lag is statistically significant which suggests that the shock effect is short-lived. Additionally, it would also be of interest to investigate whether the second lag effect differs according to the value for the first lag. However, because there are almost no subsequent failures, the interaction would hold little information. In fact, in our data there are only three cases for which the two previous patients deceased.

4.4. Transmission of Risks

The LPM's in Table VII below (columns one and three) show evidence of a transmission of shocks between adverse events. Firstly, we find that when a patient suffers a stroke during a hospitalization, (s)he is also more likely to die. As such having a stroke is strongly correlated with a procedural failure. Secondly, given that the previous patient had a stroke, the probability of mortality is lower for the next patient. The same intuition holds for the effect of a previous mortality on a subsequent stroke. However, when correcting for incidental parameters with the bias correction of De Vos *et al.* (2015), the results become insignificant and thus no longer indicate spill-over effects from one adverse event to another.

--- Insert Table VII here ---

5. ROBUSTNESS

To further test for the robustness of our results, we provide several additional tests and specifications throughout this section. Firstly, we test for the robustness of non-stationarity of regressors in the split-panel jackknife estimation. Secondly, the learning-related regressors are likely to have a non-linear relationship with patient health outcomes. We show that adding learning variables in different specifications does not qualitatively alter our results.

5.1. Irregular spacing

To account for the irregular spacing in our sample and similarly to McKenzie (2001), the dependence between observations in our pseudo-panel data set could be modelled in a linear probability model set-up as an AR(1)-process as follows:

$$Outcome_{h,t_j} = \beta outcome_{h,t_{j-1}} + u_{h,t_j} \quad (3)$$

Where:

$$u_{h,t_j} = \mu_h + v_{h,t_j}$$

For which $\mu_h \sim i.i.d(0, \sigma_\mu^2)$ and $v_{h,t_j} \sim i.i.d(0, \sigma_u^2)$. Writing our dynamic panel estimation taking into account observed time then results in the following:

$$Outcome_{h,t_j} = \beta^{t_j - t_{j-1}} outcome_{h,t_{j-1}} + \sum_{w=0}^{t_j - t_{j-1} - 1} \beta^w u_{h,t_j - w} \quad (4)$$

McKenzie (2001) shows that within this setup, the fixed effects may not be fully accounted for if the irregular spacing is not taken into account. Because of the nonlinear parameter restrictions in equation (4), there is a need to estimate with nonlinear least squares estimation. Including the intercepts, the estimation equation becomes:

$$Outcome_{h,t_j} = \beta^{t_j - t_{j-1}} outcome_{h,t_{j-1}} + \sum_{w=0}^{t_j - t_{j-1} - 1} \alpha_h \beta^w u_{h,t_j - w} \quad (5)$$

The obvious drawback of this approach is that the effect is, by assumption, highly (multiplicatively) dependent on time. The time difference in the powers are the days between procedures for our application. The specification in equation (5) then implies that the effect of a previous failure ($outcome_{h,t,j-1}$) vastly decreases over time.¹⁰ This may not be sensible for our application. The outcome of the previous patient may be important, and additive instead of multiplicative, irrespective of the time that has passed between patients. A second drawback to the estimation procedure proposed by McKenzie (2001) is the requirement that all variables are observed in all

¹⁰ E.g. suppose β equals -0.8, meaning the probability of dying after a failure is 80%-points lower. This effects decreases to only 21% points after one week.

periods, including the unobserved periods. Hence, the estimation prohibits using time-varying individual variables. This may not pose a problem in most applications because of the prominence of the lagged effect. However, we have to cope with very different patients over time. Hence, although potentially preferred in theory because of the irregular spacing of patients, the best available alternative method also suffers from substantial shortcomings.

5.2. Quadratic learning curves

As an additional robustness check, we allow for non-linearities in the relationship between the typical learning variables (cumulative experience, economies of scale and human capital depreciation) and patient mortality or having a stroke. The resulting FE LPM estimates of learning from failure effects can be found in table VIII below. Overall, the estimated effects are in line with our previous findings: previous failures are both negatively associated with short- and long-term mortality of the next patient and having complications during the procedure. Non-linearities play only a minor role in explaining these patient outcomes as all the coefficients on the squared learning variables are near to zero and therefore not economically significant.

--- Insert Table VIII here ---

6. CONCLUSION

Identifying different channels through which physicians affect patients' health may help policy makers to efficiently allocate resources to policy interventions. In this paper, we shed light on the question how procedural failures of physicians (or teams) affect subsequent patient outcomes. We show that this "learning from failure effect" is an important source of physician learning besides the commonly identified factors such as economies of scale, learning from cumulative experience

and human capital depreciation. To identify such learning from failure effects, we apply the recently developed bias-corrected fixed effects estimator by De Vos *et al.* (2015) and the split-panel jackknife estimator proposed by Dhaene and Jochmans (2015) to address the econometric challenges inherent to non-linear dynamic panel data settings.

Our findings for TAVI heart valve replacements provide evidence for a significant and sizeable negative effect from a previous failure on subsequent patient mortality. We find that a previous death is significantly associated to decrease the probability of a subsequent patient death between 6-11%-points. However, our results suggest that these effects are only short-lived, and they do not shift the slope of the cumulative learning effects. Possible short-term reasons underlying this learning from failure effect are increased concentration or short-term improved knowledge of the physicians or teams. Finally, we have illustrated the robustness of our results by showing a range of estimation techniques and employing different model specifications.

7. LITERATURE

Aakvik, A., Holmas, T.H., (2006). Access to primary health care and health outcomes: The relationships between GP characteristics and mortality rates. *Journal of Health Economics*, 25, 1139-1153.

Baltagi, B. (2005). *Econometric analysis of panel data*. John Wiley & Sons Ltd.

Bruno, G.S.F. (2005). Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals. *The Stata Journal*, 5(4), 473-500.

Bun, M. J., & Kiviet, J. F. (2006). The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. *Journal of econometrics*, 132(2), 409-444.

Bun, M. J., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *The Econometrics Journal*, 13(1), 95-126.

Cervero, R.M., Gaines, J.K. (2015). The impact of CME on physician performance and patient health outcomes: an updates synthesis of systematic reviews. *Journal of Continuing Education in the Health Professions*, 35(2),131-138.

Chernozhukov, V., Fernández-Val, I., Hahn, J., Newey, W. (2013). Average and quantile effects in nonseparable panel models. *Econometrica*, 81(2), 535-580.

Cruz-Gonzalea, M., Fernández-Val, I., Weidner, M. (2017). Bias corrections for probit and logit models with two-way fixed effects. *The Stata Journal* 17(3):517-545.

- De Vos, I., Everaert, G., & Ruysen, I. (2015). Bootstrap-based bias correction and inference for dynamic panels with fixed effects. *Stata Journal (forthcoming)*, 1-31.
- Dhaene, G., Jochmans, K. (2015). Split-panel jackknife estimation of fixed-effect models. *Review of Economics Studies*, 82, 991-1030.
- Edmondson, A.C. (2004). Learning from failure in health care: frequent opportunities, pervasive barriers. *Quality and Safety in Health Care*, 13(Sup. II), ii3-ii9.
- Eggebrecht, H., Schäfer, U., Treede, H., Boekstegers, P., Babin-Ebell, J. et al. 2011. Valve-in-Valve Transcatheter Aortic Valve Implantation for Degenerated Bioprosthetic Heart Valves. *JACC: Cardiovascular Interventions* **4(11)**: 1218-1227.
- Everaert, G., Pozzi, L. (2007). Bootstrap-based bias correction for dynamic panels. *Journal of Economic Dynamics & Control*, 31, 1160-1184.
- Fernández-Val, I., Weidner, M. (2016). Individual and time effects in nonlinear panel models with large N, T. *Journal of Econometrics*, 192, 291-312.
- Gaynor M., Seider H., Vogt W.B. (2005). The Volume-Outcome Effect, Scale Economies, and Learning-by-Doing. *The American Economic Review: Papers and Proceedings*, **95(2)**, 243-247.
- Ghanem, A., Naderi, A.S., Frerker, C., Nickenig, G., Kuck, K-H. 2016. Mechanisms and Prevention of TAVI-related Cerebrovascular Events. *Current Pharmaceutical Design* **22**: 1-9.
- Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal*, 7(1), 98-119.

Hareli, S., Karnieli-Miller, O., Hermoni, D., Eidelman, S. (2007). Factors in the doctor-patient relationship that accentuate physicians' hurt feelings when patients terminate the relationship with them. *Patient Education and Counseling*, 67, 169-175.

Hentschker, C., Mennicken, R. (2015), The Volume-outcome Relationship and Minimum Volume Standards - Empirical Evidence for Germany. *Health Economics*, 24(6), 644-658.

Hockenberry, JM., Helmchen L.A., (2014). The nature of surgeon human capital depreciation. *Health Economics*, 37, 70-80.

Iizuka, T., Watanabe, Y., (2016). The impact of physician supply on the healthcare system: evidence from Japan's new residency program. *Health Economics*, 25, 1433-1447.

Janakiraman, R., Dutta, S., Sismeiro, C., Stern, P. (2008). Physicians' persistence and its implications for their response to promotion of prescription drugs. *Management Science*, 54(6), 1080-1093.

Jollis, J.G., DeLong, E.R., Peterson, E.D., Muhlbaier, L.H., Fortin, D.F., Califf, R.M. Mark, D.B. (1996). Outcome of acute myocardial infarction according to the specialty of the admitting physician. *The New England Journal of Medicine*, 335, 1880-1887.

Jones, N.P., Papadakis, A.A., Orr, C.A., Strauman, T.J. (2013). Cognitive processes in response to goal failure: a study of ruminative thought and its affective consequences. *Journal of Social and Clinical Psychology*, 23(5), 1-19.

Kolstad, J.T (2013), Information and Quality when Motivation is Intrinsic: Evidence from Surgeon Report Cards. *American Economic Review*, 103(7), 2875-2910.

- Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of econometrics*, 95(2), 391-413.
- Li, J., Hurley, J. Decicca, P., Buckley, G. (2014). Physician response to pay-for-performance: evidence from a natural experiment. *Health Economics*, 23, 963-978.
- Liu, R. Y. (1988). Bootstrap procedures under some non-iid models. *The Annals of Statistics*, 16(4), 1696-1708.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The annals of statistics*, 255-285.
- Moon, H.R., Perron, B., Phillips, P.C.B. (2015). Incidental parameters and dynamic panel modelling in “The Oxford Handbook of Panel Data” (ed. Badi H. Baltagi)
- McKenzie, D. J. (2001). Estimation of AR (1) models with unequally spaced pseudo-panels. *The Econometrics Journal*, 4(1), 89-108.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417-1426.
- Oberlander, J. (2007). Learning from Failure in Health Care Reform. *The New England Journal of Medicine*, 357(17), 1677-1679.
- Or, Z., Wang, J., Jamison, D. 2005. International differences in the impact of doctors on health: a multilevel analysis of OECD countries. *Journal of Health Economics*, 24,531-560.
- Panthöfer, S. (2017). Do Doctors Prescribe Antibiotics Out of Fear of Malpractice? Mimeo.

Pua, A. A. Y. (2015). On IV estimation of a dynamic linear probability model with fixed effects (No. 15-01). Universiteit van Amsterdam, Department of Econometrics.

Rizzo, J.A., Zeckhauser, R.J. (2003). Reference incomes, loss aversion, and physician behavior. *The Review of Economics and Statistics*, 85(4), 909-922.

Salge, T.O., Vera, A., Antons, D., Cimiotti, J.P. (2016). Fighting MRSA infections in hospital care: how organizational Factors Matter. *Health Services Research*, 52(3), 959-983.

Shepherd, D.A., Haynie, M., Patzelt, H. (2013). Project failures arising from corporate entrepreneurship: impact of multiple project failures on employees' accumulated emotions, learning, and motivation. *Journal of Product Innovation Management*, 30(5), 880-895.

Sundmacher, L., Busse, R., (2011). The impact of physician supply on avoidable cancer deaths in Germany. A spatial analysis. *Health Policy*, 103, 53-62.

Van Gestel, R., Müller, T., Bosmans, J. Forthcoming. Does My High Blood Pressure Improve Your Survival? Overall and Subgroup Learning in Health. *Health Economics*.

Ward, M.M., Yankey, J.W., Vaughn, T.E., BootsMiller, B.J., Flach, S.D., Welke, K.F., Pendergast, J.F., Perlin, J., Doebbeling, B.N. (2004). Physician Process and Patient Outcome Measures for Diabetes care: Relationships to Organizational Characteristics. *Medical Care*, 42(9), 840-850.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

APPENDIX A: A MONTE CARLO SIMULATION OF DYNAMIC PANEL BIAS-CORRECTION MODELS

In this section, we shed light on the relative performance of the bias-corrected FE estimator proposed by De Vos *et al.* (2015) in the context of dynamic binary outcome variables. Specifically, we run Monte Carlo simulations to illustrate the finite sample properties of the fixed effects LPM (FE LPM) and the bias-corrected FE (BCFE) estimator. This exercise is informative as De Vos *et al.* (2015) only present the relative performance of their bias-correction method in case of continuous outcome variables. In addition, the simulation takes into account the irregular spacing of observations to assess the adequacy of the proposed methods in this setting.

Our simulation results are based on the following data generating process:

$$(1) \quad y_{it} = \alpha_i + 0.5 y_{it-1} + \varepsilon_{it}$$

$$(2) \quad \varepsilon_{it} = \lambda_i F_t + v_{it}$$

where y_{it} is a latent continuous dependent variable, $\alpha_i \sim iid \mathcal{N}(0, \varphi)$ is an unobserved individual effect and ε_{it} is a cross-sectionally dependent error term with $\lambda_i \sim iid U(1,4)$, $F_t \sim iid \mathcal{N}(0,1)$ and $v_{it} \sim iid \mathcal{N}(0,1)$. First, we explicitly allow for cross-sectional dependence in the errors. In our application this implies that across hospitals, similar patients (corrected for observed covariates) are treated at certain experience levels¹¹. Initial values for the dependent variable are randomly drawn from a standard normal. We subsequently dichotomize the latent dependent variable based on the following threshold mechanism: $y_{it}^{observed} = I(y_{it} \geq 0)$. To mimic the irregularly spaced TAVI data set, we generate different panel lengths ranging from $T = 3-5$ for the different

¹¹ Remember that our panel time dimension is based on experience.

observations. Moreover, we use $N = 800$ observations in each simulation sample to again closely resemble the data dimension in our actual data set used in the paper.

Figure A1 below shows the result from our Monte Carlo experiment. Our results suggest that FE LPM estimates suffer from substantial finite sample biases as they are nowhere close to the true ADPE. On the other hand, the BCFE estimator is substantially less biased and clearly outperforms the FE LPM¹².

Second, as an additional exercise, we check the relative performance of the two estimators in case of autocorrelated errors. This means in our application that error terms for different patients are correlated over time within a hospital, while including patient characteristics and hospital fixed effects. This could happen when providers select on unobserved characteristics in a time-dependent manner. Error terms were generated based on the following AR(1)-process:

$$(3) \quad \varepsilon_{it} = \gamma + \phi \varepsilon_{it-1} + v_{it}$$

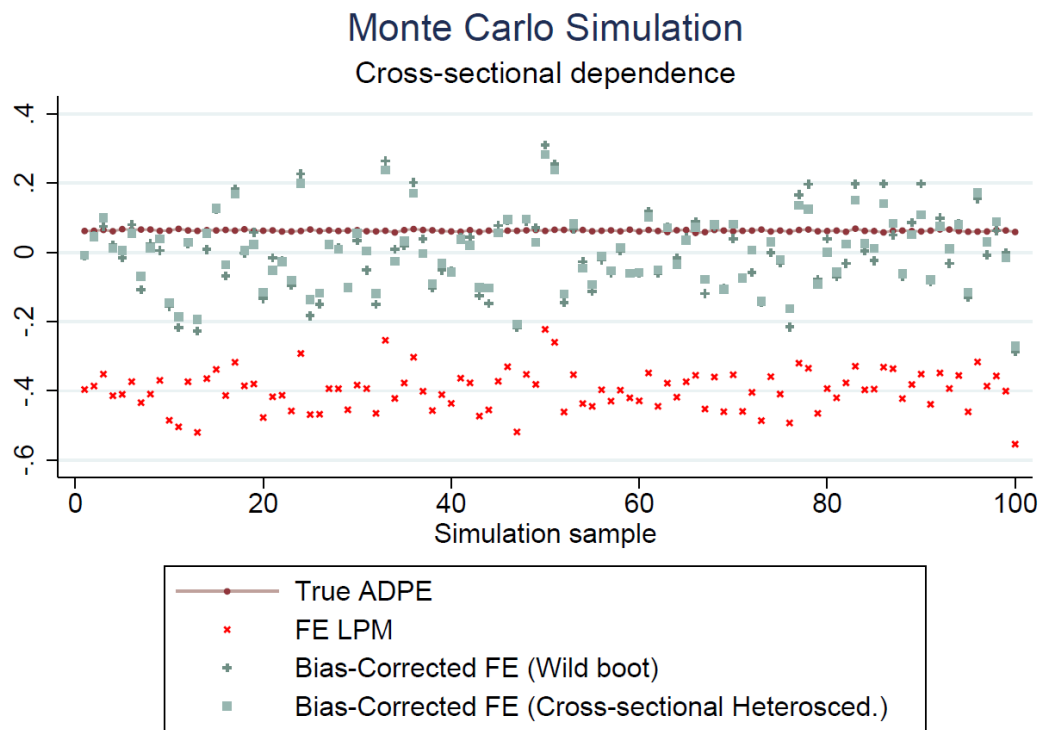
Initial values for ε_{it} were drawn from an iid standard normal and we simulate based on two parametrizations for ϕ : in the first setup, we implement a negative autocorrelation in the error of $\phi = -0.6$ and in the second, a positive one with $\phi = +0.6$.

The corresponding Monte Carlo experiment shows that the introduction of serial correlation in general induces quite substantial distortions onto the estimated effects. This suggests that are our results are not fully robust to autocorrelated errors, which may exist when providers select patients on unobserved characteristics and when this selection is time-dependent. Although certainly an

¹² We have also run a similar Monte Carlo experiment with iid errors and the results are even more in favor of the BCFE estimator than in the above case.

interesting finding, the corresponding econometric issue would have to be carefully analyzed theoretically and empirically in a stand-alone future research project¹³.

Figure A1: Monte Carlo simulation results – Cross-sectional dependence



¹³ The corresponding simulation results are available upon request.

TABLES AND FIGURES

Table I: Descriptive Statistics

Variable	N	Mean	Std. Dev.	Variable	N	Mean	Std. Dev.
One-month mortality	854	0.093	0.290	Type Valve	855	0.464	0.499
24-month mortality	854	0.285	0.451	Transfemoral Access	852	0.742	0.438
Angina	841	0.278	0.448	Size valve 1	838	0.520	0.500
Aortic Aneurysm	841	0.057	0.232	Size valve 2	838	0.243	0.429
Atrial Fibrillation	839	0.300	0.459	Size valve 3	838	0.014	0.119
Carotid Disease	835	0.196	0.398	<u>Total number of TAVI procedures</u>			
Coronary Artery Disease	840	0.606	0.489	Hospital 1	77		
COPD	842	0.287	0.453	Hospital 2	30		
Chronic Heart Failure	837	0.665	0.472	Hospital 3	19		
Diabetes	841	0.251	0.434	Hospital 4	13		
Hypertension	840	0.749	0.434	Hospital 5	18		
Pulmonary Hypertension	837	0.548	0.498	Hospital 6	25		
Myocard Infarct	839	0.225	0.418	Hospital 7	24		
Porcelain Aorta	841	0.087	0.282	Hospital 8	20		
Renal Failure	840	0.251	0.434	Hospital 9	34		
PCI	841	0.310	0.463	Hospital 10	19		
Pacemaker	849	0.133	0.340	Hospital 11	73		
Stroke TIA	839	0.153	0.360	Hospital 12	74		
Ejection Fraction	821	53.414	14.773	Hospital 13	15		
NYHA category 2	839	0.180	0.384	Hospital 14	28		
NYHA category 3	839	0.604	0.489	Hospital 15	32		
NYHA category 4	839	0.200	0.400	Hospital 16	15		
Mediastinal Radiation	839	0.043	0.203	Hospital 17	76		
Defibrillation	840	0.004	0.060	Hospital 18	25		
CABG	839	0.263	0.441	Hospital 19	126		
Valve Surgery	840	0.030	0.170	Hospital 20	112		

The table shows the summary statistics for the observed outcomes, patient and procedure characteristics, as well as the total number of TAVI procedures per hospital over the years 2007-2012.

Table II: Descriptive Statistics - Patients undergoing TAVI in Belgium

	2007	2008	2009	2010	2011	2012	Total
Total	10	100	163	257	289	36	855

The table shows the total number of patients undergoing the TAVI procedure in Belgium over the sample period from 2007-2012.

Table III: Adverse event after a non-failure and after a failure

	1-month mortality		Stroke		Renal Failure		Pacemaker	
	Lag=0	Lag=1	Lag=0	Lag=1	Lag=0	Lag=1	Lag=0	Lag=1
Number of Adv. Events	75	3	27	0	81	15	87	22
Total number of individuals	756	78	711	30	641	99	623	113
Prob. Adv. Event	0.099	0.038	0.038	0.000	0.126	0.152	0.140	0.195
P-val.	0.013		0		0.514		0.168	

In columns Lag=0, the probability of the adverse events provided for individuals for which the previous patient did not encounter the adverse event. In the Lag=1 column, the previous patients did encounter the adverse event. The p-value refers to the test for equality of proportions between the Lag=0 and Lag=1 columns.

Table IV: FE LPM & FE Logit estimates of the learning from failure effects

Outcome	1-month mortality FE LPM	1-month mortality FE Logit ADPE	24-month mortality FE LPM	24-month mortality FE Logit ADPE
<i>Lagged Outcome</i>	-0.078*** (0.025)	-0.081*** (0.021)	-0.074** (0.036)	-0.075** (0.036)
<i>CE</i>	0.000 (0.000)	0.000 (0.001)	-0.002** (0.001)	-0.002** (0.001)
<i>EOS</i>	-0.001 (0.021)	-0.001 (0.002)	0.005 (0.003)	0.006 (0.004)
<i>HCD</i>	0.000 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)
Controls	Yes	Yes	Yes	Yes
Hospital FE's	Yes	Yes	Yes	Yes
Observations	762	731	762	762

FE LPM and FE Logit average discrete probability effects (ADPE) are shown for the learning from failure effects for the outcomes of one- and 24-month mortality. The average baseline probabilities for one- month mortality is 8.8% and 28.9% for 24-month mortality. Heteroscedasticity-robust standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

Table V: Bias-corrected FE LPM estimates of the learning from failure effects

Outcome	1-month mortality	24-month mortality
<i>Lagged outcome</i>	-0.046 (0.033)	-0.060* (0.033)
<i>CE</i>	-0.001 (0.001)	-0.003*** (0.001)
<i>ESC</i>	0.007** (0.003)	0.008* (0.005)
<i>HCD</i>	0.000 (0.000)	0.000 (0.001)
Controls	Yes	Yes
Hospital FE	Yes	Yes
Observations	550	550

Learning from failure effects are estimated using the bias-corrected FE LPM proposed by De-Vos et al. (2015) for the outcomes of one- and 24-month mortality. Bootstrapped standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

Table VI: Split-panel jackknife FE probit estimates of the learning from failure effects¹⁴

Outcome	1-m mortality	24-m mortality
<i>Lagged outcome</i>	-0.112*** (0.020)	-0.070** (0.034)
<i>CE</i>	0.001 (0.001)	0.000 (0.001)
<i>ESC</i>	-0.005 (0.002)	-0.001 (0.003)
<i>HCD</i>	- -	-0.000 (0.001)
Controls	Yes	Yes
Hospital FE	Yes	Yes
Observations	730	761

The table shows the estimates of the learning from failure effects using the Split-panel jackknife FE probit estimator proposed by Dhaene and Jochmans (2015). Standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1. For 1-month mortality, the estimates do not converge when HCD is included. Overall, the coefficient of the Lagged outcome is robust to leaving out the learning variables (i.e., leaving out all learning variables for 1-month mortality and leaving out HCD for 2-year mortality). Restricting the sample size to 730 instead of 761 for 24-m mortality does not change the results dramatically (the point estimate is -0.062 and statistically significant at the 10% level).

¹⁴ Note that we use the *probitfe* command here, we tested robustness with the *logitfe* command. Using *logitfe*, the log likelihood function does not converge.

Table VII: Transmission of shocks

Outcome	Mortality LPM	Mortality Bias Corr.	Stroke LPM	Stroke Bias Corr.
Lagged Mortality	-0.052*	-0.003	-0.024**	-0.018
	(0.028)	(0.056)	(0.012)	(0.042)
Stroke	0.240***	0.218***	-	-
	(0.091)	(0.059)	-	-
Lagged Stroke	-0.047*	-0.053	-0.053**	-0.059
	(0.026)	(0.057)	(0.025)	(0.050)
Controls	Yes	Yes	Yes	Yes
Observations	671	479	671	477
R-squared	0.137	-	0.107	-

The Table shows the estimated spill-over effects from one adverse event to others using a simple FE LPM (columns 1 and 3) and bias-corrected FE LPM (columns 2 and 4) estimators. Robust standard errors in parentheses in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

Table VIII: Inclusion of squared learning variables

Outcomes	1-m mortality	24-m mortality
Lagged outcome	-0.080*** (0.025)	-0.076** (0.036)
CE	0.002 (0.002)	-0.001 (0.003)
ESC	0.010 (0.007)	0.013 (0.010)
HCD	0.000 (0.001)	-0.000 (0.001)
CE squared	-0.000 (0.000)	-0.000 (0.000)
ESC squared	-0.000* (0.000)	-0.000 (0.000)
HCD squared	-0.000 (0.000)	0.000 (0.000)
Same day	0.047 (0.030)	0.021 (0.047)
Controls	Yes	Yes
Hospital FE	Yes	Yes
Observations	762	762
R-squared	0.098	0.117

The table shows the FE LPM estimates of the learning from failure effects including quadratic terms of the learning variables (cumulative experience, economies of scale and human capital depreciation). Heteroscedasticity-robust standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1

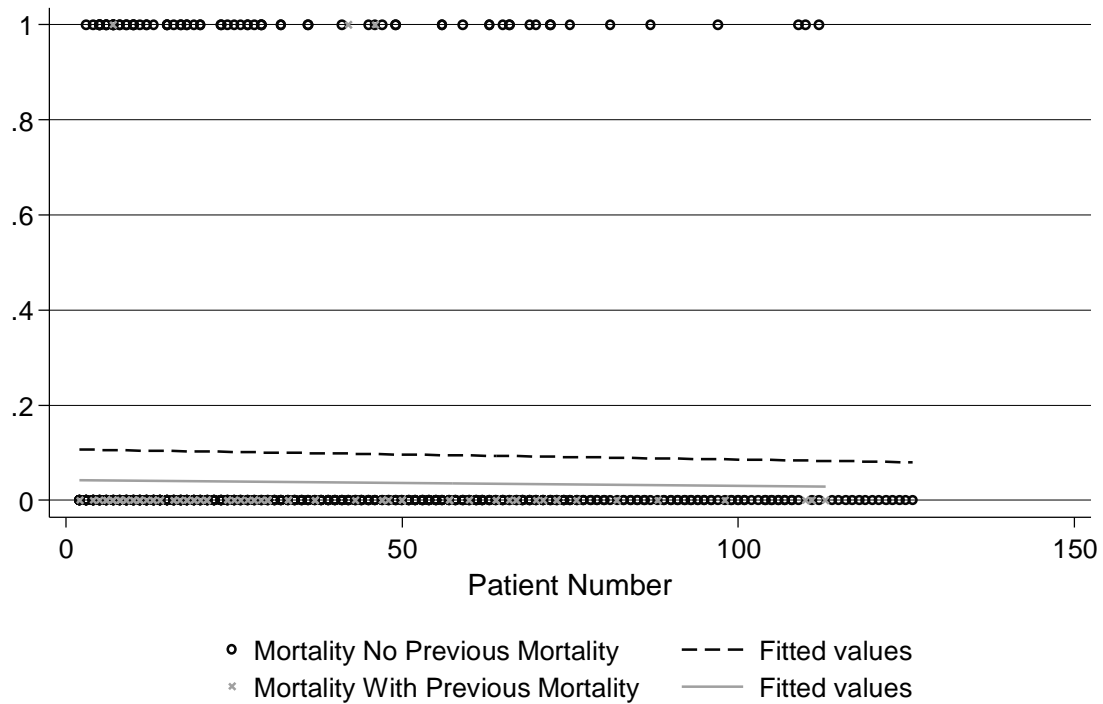


Figure 1: Learning pattern with and without previous failure

The dots represent individual data on mortality (0=did not die, 1=did die), conditional on the value for mortality of the previous patient (black= the previous patient survived, grey=the previous patient died). The fitted lines represent the effect of experience (patient number) on mortality conditional on whether the previous patient died or not (grey or black respectively).