Impacts of Care Provider Collaborations on the Service Time for Inpatient Stays: An Analysis using EHR Audit Logs and Dynamic Graphs

Gaurav Jetley

Computer Information Systems, College of Business, Colorado State University, Fort Collins, CO

He Zhang

Department of Technology & Operations Management, College of Business Administration, California State Polytechnic University, Pomona, CA

Appendix A – Illustration of Captured Collaborations using Collaboration Network

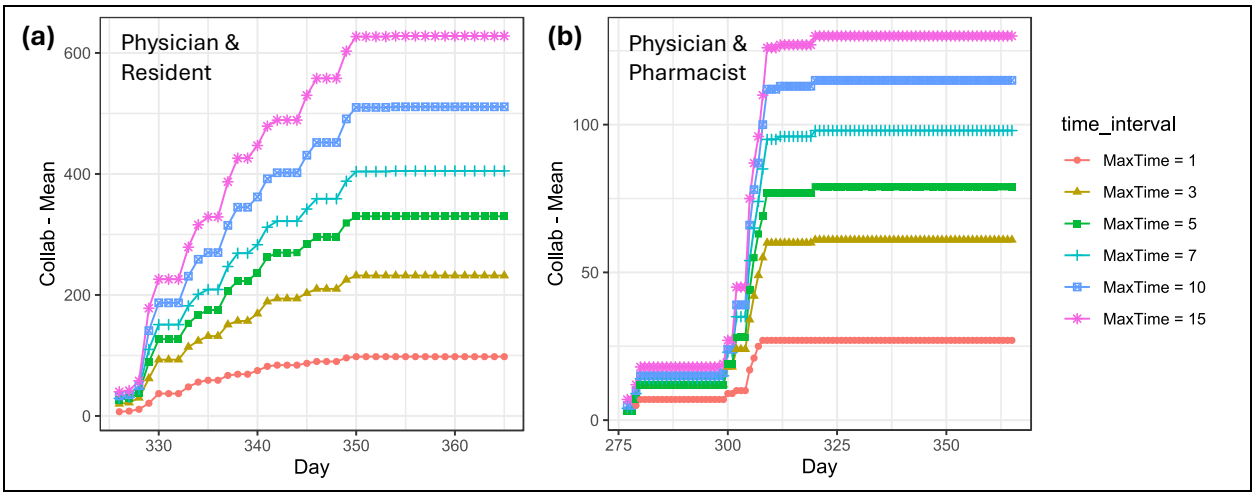

Figure A.1: Illustration of Captured Collaborations

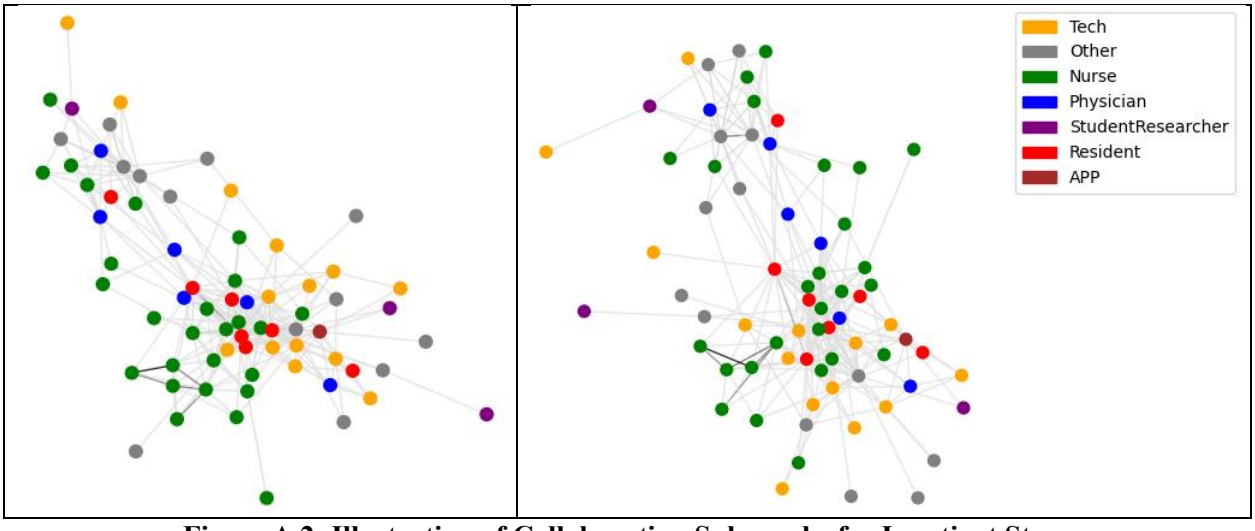

Figure A.2: Illustration of Collaboration Subgraphs for Inpatient Stays

Table A.1.: Distinguishing Familiarity vs. Collaboration

Familiarity	Collaboration	Description	Real-World Scenario
High	**High**	The lead physician and the care team member have worked together across many patient stays and frequently access the same patient record around the same time.	A lead hospitalist and a case manager who are both regularly assigned to the same ward. They typically review and update charts during daily rounds and discharge planning meetings.
High	**Low**	The two providers appear on many of the same patient stays but rarely access the EHR near the same time.	A physician and a unit nurse who both cover the same set of patients across multiple admissions—but with staggered schedules, resulting in minimal real-time interaction.
Low	**High**	The providers rarely overlap across stays but jointly access the record in short time windows—possibly due to a high-acuity event.	A night-call physician collaborating closely with an ICU pharmacist during an emergency medication review for a deteriorating patient. They've never worked together before, but both access the chart intensively within minutes.
Low	**Low**	The providers have little historical overlap and show no temporally proximate EHR access.	A consulting nephrologist and a speech therapist both involved in a patient's care plan at different points, with no shared work history or time-aligned engagement.

Appendix B – Estimation Strategy

B.1. First Stage Estimates and *F*-test for Excluded Instruments

The *F*-test for excluded instruments results for the 6 models are provided below. We follow the rule of thumb of this test statistic being greater than 10 for strong instruments. We used *ivreg2* user written command from Stata to run the IV regression models and obtain the *F*-test results.

Table B.1.1.: *F*-Test for Excluded Instruments and Weak Identification Test for Max_{Time} = 1, 3, 5, 7, 10 & 15

	***F*-Test of Excluded Instruments**			**Sanderson-Windmeijer *F*-Test**		
Model	**Collab–Mean**	**(Collab–Mean)^2**	**Collab–SD**	**Collab–Mean**	**(Collab–Mean)^2**	**Collab–SD**
$Max_{Time} = 1$	**62.03*** **	**52.42*** **	**56.3*** **	**38.81*** **	**26.46*** **	**90.81*** **
$Max_{Time} = 3$	**102.39*** **	**96.41*** **	**61.09*** **	**63.43*** **	**56.86*** **	**57.71*** **
$Max_{Time} = 5$	**143.86*** **	**173.58*** **	**61.32*** **	**42.09*** **	**37.68*** **	**61.99*** **
$Max_{Time} = 7$	**144.96*** **	**160.58*** **	**69.46*** **	**33.67*** **	**30.76*** **	**105.05*** **
$Max_{Time} = 10$	**151.53*** **	**155.07*** **	**76.00*** **	**50.06*** **	**46.81*** **	**139.57*** **
$Max_{Time} = 15$	**175.81*** **	**169.65*** **	**89.58*** **	**115.44*** **	**108.03*** **	**150.73*** **

B.2. Placebo Test

Table B.2.1: Placebo Test

	(1)	**(2)**	**(3)**
	$Max_{Time} = 1$	$Max_{Time} = 1$	$Max_{Time} = 1$
	OLS	**OLS**	**OLS**
Collab – Mean	-0.155***		
	(0.036)		
Collab – Mean^{IV}	0.005		
	(0.088)		
(Collab – Mean)^2		-0.191***	
		(0.025)	
(Collab – Mean)$^{2\ IV}$		-0.044	
		(0.07)	
Collab – SD			0.013**
			(0.004)
Collab – SD^{IV}			0.012
			(0.025)
Constant	1.146***	1.107***	1.102***
	(0.151)	(0.152)	(0.163)
Observations	8,405	8,405	8,402

Appendix C – Results, Additional Analysis and Robustness Checks

C.1. Main Results for Max_{Time} = 7, 10 and 15 minutes Models

Table C.1.1.: Overall Collaborations and LOS

	(1)	(2)	(3)
	$Max_{Time} = 7$	$Max_{Time} = 10$	$Max_{Time} = 15$
	IV-2SLS	**IV-2SLS**	**IV-2SLS**
Collab – Mean	5.694*	4.656*	3.599*
	(2.474)	(2.09)	(1.679)
(Collab – Mean)^2	-2.09**	-1.583**	-1.114**
	(0.73)	(0.562)	(0.397)
Collab – SD	0.087***	0.065***	0.043***
	(0.018)	(0.013)	(0.01)
***F*-Test of Excluded Instruments:**			
	144.96***	151.53***	175.81***
	160.58***	155.07***	169.65***
	69.46***	76.00***	89.58***
Constant	-2.691	-2.341	-1.861
	(1.772)	(1.698)	(1.568)
Observations	8,406	8,406	8,406

Note:
**** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1*
Robust standard errors (clustered on CCS category) in parentheses.
e(Collab – Mean^{IV}) and e(Collab – Mean^2 ^{IV}) are the residuals from the first-stage regression which are introduced in the second-stage for the control function approach.

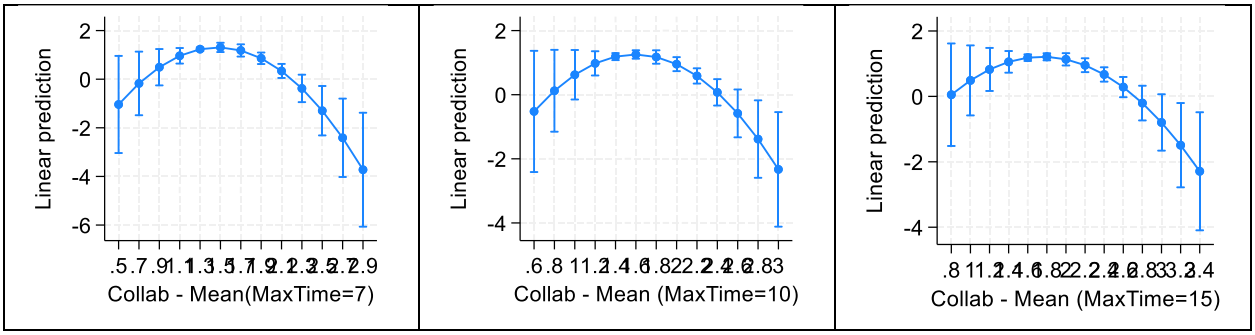

Figure C.1.2.: Marginal Effects

C.2. Daily Collaboration Analysis Using Survival Models

Survival Analysis: Estimation Strategy and Instrumental Variables

In addition to analyzing the impact of overall collaboration strength on length of stay (LOS), we also examine how daily collaborations among care providers influence the time to discharge for patients. To capture this dynamic effect, we employ parametric survival analysis, which is suitable for modeling time-to-event data and also allows for instrumental variable (IV) estimation—a feature for which the Cox proportional hazards (Cox PH) model is not suitable. We test multiple distributions, including Weibull, Exponential, Log-Normal, Log-Logistic, and Generalized Gamma. Based on the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values, we select the Generalized Gamma distribution for our parametric survival model. This model assumes a generalized gamma distribution for the survival time and captures the effect of daily collaborations on discharge. By modeling the hazard function in this way, we can assess how fluctuations in daily collaboration levels affect the probability of patient discharge over time while controlling for other influential factors.

Instrumental Variable Strategy for Daily Collaborations

Potential endogeneity arises if unobserved factors, such as patient severity not fully captured by our measures, influence both daily collaboration levels and the time to discharge. To address this concern, we implement an instrumental variable (IV) approach tailored for the daily collaboration context.

Constructing a valid IV for daily collaborations is more complex than for overall collaboration strength due to varying lengths of stay (LOS) among patients. Directly averaging daily collaborations across non-focal stays is problematic because patients differ in LOS, making daily collaboration patterns misaligned across stays. To overcome this, we standardize daily collaboration patterns by applying cubic spline interpolation to the daily collaboration data from non-focal inpatient stays under the same lead physician.

Cubic spline interpolation allows us to create smooth, continuous functions that estimate daily collaboration levels over a standardized time frame. The interpolation process is defined as:

$$\widehat{DailyCollabs}_j(t) = spline\left(DailyCollabs_j \sim t\right), t = 1,2, \dots, \max(LOS_i)$$

where $\widehat{DailyCollabs}_j(t)$ represents the interpolated collaboration values for the non-focal stay *j* over a standardized time frame matching the LOS of the focal stay *i*. This interpolation, normalizes collaboration patterns across different LOS, making them comparable.

For each focal patient stay *i*, we construct the IV for daily collaboration as the average of the interpolated daily collaborations from the non-focal stays under the same lead physician and did not overlap with the focal stay (i.e., discharge before or admission after the focal stay):

$$DailyCollabsIV_i(t) = \frac{1}{N_i} \sum_{j \in J_i} \widehat{DailyCollabs}_j(t)$$

where $N_i = |J_i|$ is the number of non-focal stays under lead physician *p*, and J_i denotes the set of these non-focal stays. This average represents the exogenous component of daily collaboration that is informed by collaboration patterns external to the focal stay *i*. This IV represents the typical daily collaboration pattern associated with the lead physician, independent of the focal patient's specific characteristics or unobserved severity.

For the IV to be valid, it must satisfy two key conditions: relevance and exogeneity. First, the IV must be correlated with the endogenous variable (daily collaboration strength). Lead physicians often exhibit consistent collaboration patterns due to personal practice styles, habitual interactions, and established communication networks. Empirically, we find a significant positive correlation between $DailyCollabsIV_i(t)$ and the endogenous variable $DailyCollabs_i(t)$, confirming that the IV is strongly associated with the daily collaboration levels of the focal patient stay. We also find the *F*-test of excluded instruments to be above 10, which tells us that the IV is strong and relevant.

Second, the IV must not be correlated with the error term in the outcome equation. The IV is constructed from non-focal stays, which are independent of the unobserved patient-specific factors affecting LOS for the focal stay. While the lead physician's collaboration style influences both the IV and the focal stay's collaboration, it does not directly affect the focal patient's LOS except through collaboration. Staffing decisions, care plans, and treatment procedures are tailored to individual patient needs, minimizing the

potential for direct influence from non-focal collaboration patterns. Therefore, the IV affects LOS only indirectly via its impact on daily collaboration levels.

We use a control function approach which is appropriate for nonlinear models like the parametric survival model. We did not use the Cox Proportional Hazards model because the two-stage residual inclusion (2SRI) method can introduce bias in semi-parametric models like the Cox model.

In the first stage, we regress the endogenous variable on the IV and other exogenous covariates:

$$\log\big(DailyCollabs_i(t)\big) = X_i\Pi + \pi_1 DailyCollabsIV_i(t) + \upsilon_i(t)$$

where Π and π_1 are coefficients to be estimated, and $\upsilon_i(t)$ is the error term capturing unobserved factors. In the second stage, we include the residuals from the first stage in the survival model to correct for endogeneity. This approach allows us to capture the dynamic nature of collaborations in inpatient settings and their direct impact on patient outcomes, while mitigating biases due to unobserved confounding factors. By standardizing daily collaboration patterns and employing the 2SRI method, we enhance the validity of our estimates regarding the effect of daily collaborations on the time to discharge.

Table C.2.1.: IV-Survival Analysis - Daily Collaborations and Time to Discharge

	(1)	(2)	(3)	(4)
	$\boldsymbol{Max_{Time} = 1}$	$\boldsymbol{Max_{Time} = 1}$	$\boldsymbol{Max_{Time} = 3}$	$\boldsymbol{Max_{Time} = 5}$
	IV-Survival	**IV-Survival**	**IV-Survival**	**IV-Survival**
Collab – Mean (Daily)	0.133	-0.119***	-0.149***	-0.164***
	(0.091)	(0.015)	(0.011)	(0.015)
(Collab – Mean)^2 (Daily)	-0.228**			
	(0.068)			
Collab – SD (Daily)	0.055***	0.027***	0.014***	0.01***
	(0.007)	(0.003)	(0.001)	(0.001)
e(Collab – Mean^{IV}) (Daily)	0.018	0.129***	0.146***	0.162***
	(0.053)	(0.012)	(0.027)	(0.037)
e(Collab – Mean$^{2\,IV}$) (Daily)	0.097*			
	(0.042)			
e(Collab – SD^{IV}) (Daily)	-0.034***	-0.024***	-0.012**	-0.009***
	(0.007)	(0.004)	(0.004)	(0.003)
Constant	1.362***	1.392***	1.452***	1.504***
	(0.037)	(0.032)	(0.033)	(0.036)
Observations	31,717	31,717	31,717	31,717
Subjects	6,814	6,814	6,814	6,814
Note: *** $p<0.001$, ** $p<0.01$, * $p<0.05$, ^ $p<0.1$ *Robust standard errors (clustered on CCS category) in parentheses.*				

Table C.2.2.: IV-Survival Analysis with Time Interaction

	(1)	(2)	(3)
	$Max_{Time} = 1$	$Max_{Time} = 3$	$Max_{Time} = 5$
	IV-Survival	**IV-Survival**	**IV-Survival**
Collab – Mean (Daily)	-0.056	-0.104***	-0.133***
	(0.035)	(0.027)	(0.022)
Collab – Mean (Daily) * t	-0.013*	-0.003	0.002
	(0.005)	(0.004)	(0.003)
Collab – SD (Daily)	0.019***	0.008***	0.005***
	(0.003)	(0.002)	(0.001)
e(Collab – Mean^{IV}) (Daily)	0.084**	0.077*	0.082**
	(0.026)	(0.032)	(0.034)
e(Collab – Mean^{IV} * t) (Daily)	0.006**	0.009^	0.007
	(0.004)	(0.005)	(0.005)
e(Collab – SD^{IV}) (Daily)	-0.02***	-0.009***	-0.005*
	(0.003)	(0.002)	(0.001)
Constant	0.885***	0.966***	1.036***
	(0.074)	(0.078)	(0.075)
Observations	31,717	31,717	31,717
Subjects	6,814	6,814	6,814

Note:
*** *p<0.001,* ** *p<0.01,* * *p<0.05,* ^ *p<0.1*
Robust standard errors (clustered on CCS category) in parentheses.

C.3. Role Level Collaborations

Table C.3.1.: Role Level Overall Collaborations and LOS

	(1)	(2)	(3)
	$Max_{Time} = 1$	$Max_{Time} = 1$	$Max_{Time} = 1$
	OLS	OLS	OLS
Collab – Mean (Physician – Resident)	0.266***	0.198***	0.157***
	(0.025)	(0.026)	(0.046)
(Collab – Mean)^2 (Physician – Resident)	-0.378***	-0.339***	-0.334***
	(0.017)	(0.016)	(0.021)
Collab – SD (Physician – Resident)	0.093***	0.091***	0.089***
	(0.006)	(0.006)	(0.009)
Collab – Mean (Physician – Nurse)		0.062	0.171**
		(0.047)	(0.05)
(Collab – Mean)^2 (Physician – Nurse)		-0.876***	-0.956***
		(0.058)	(0.09)
Collab – SD (Physician – Nurse)		0.253***	0.232***
		(0.019)	(0.025)
Collab – Mean (Physician – Physician)			0.037*
			(0.014)
(Collab – Mean)^2 (Physician – Physician)			-0.219***
			(0.017)
Collab – SD (Physician – Physician)			0.075***
			(0.01)
Constant	1.77***	1.757***	2.005***
	(0.049)	(0.059)	(0.087)
Observations	5,344	5,048	2,837

Note:
*** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1
Robust standard errors (clustered on CCS category) in parentheses.

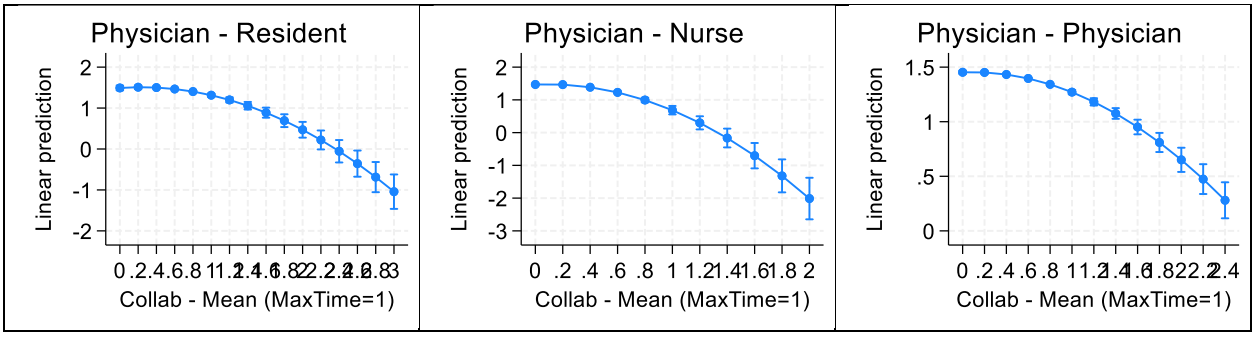

Figure C.3.1: Marginal Effects for Role Level Overall Collaborations on LOS

C.4. Differential Tipping Points in Collaboration: Surgical vs. Non-Surgical Care

Table C.4.1.: Overall Collaborations and LOS for Surgical Services

	(1)	(2)	(3)
	$Max_{Time} = 1$	$Max_{Time} = 3$	$Max_{Time} = 5$
	OLS	**OLS**	**OLS**
Collab – Mean	1.398***	0.83***	0.863**
	(0.298)	(0.181)	(0.227)
(Collab – Mean)^2	-2.01***	-0.974***	-0.803***
	(0.179)	(0.099)	(0.121)
Collab – SD	0.32***	0.189***	0.138***
	(0.032)	(0.026)	(0.024)
Constant	0.277	0.449**	0.408*
	(0.172)	(0.131)	(0.184)
Observations	1,524	1,524	1,524

Note:
**** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1*
Robust standard errors (clustered on CCS category) in parentheses.
e(Collab – Mean^{IV}) and e(Collab – Mean$^{2\,IV}$) are the residuals from the first-stage regression which are introduced in the second-stage for the control function approach.

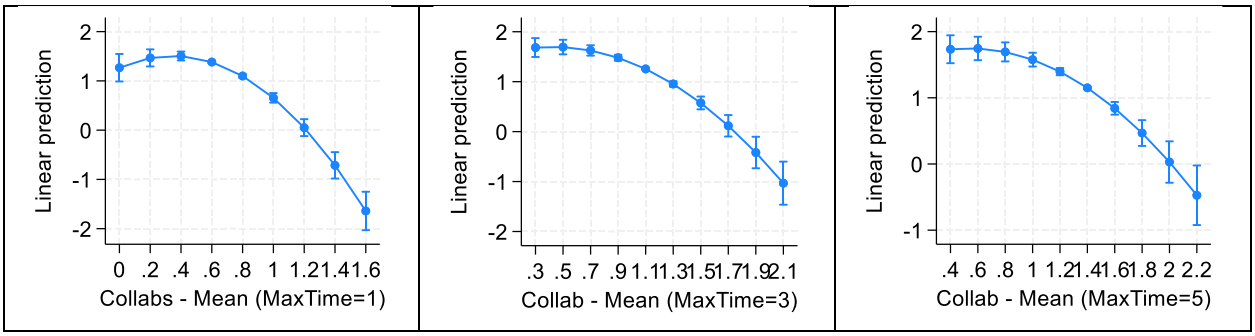

Figure C.4.1.: Marginal Effects for Surgical Services

Table C.4.1.: Overall Collaborations and LOS for Non-Surgical Services

	(1)	(2)	(3)
	$Max_{Time} = 1$	$Max_{Time} = 3$	$Max_{Time} = 5$
	IV-2SLS	**IV-2SLS**	**IV-2SLS**
Collab – Mean	7.554*	3.85	2.935
	(3.709)	(2.675)	(2.155)
(Collab – Mean)^2	-5.988***	-2.047*	-1.374*
	(2.331)	(1.003)	(0.677)
Collab – SD	0.25**	0.105*	0.076*
	(0.084)	(0.048)	(0.03)
***F*-test of Excluded Instruments**	64.40***	115.21***	97.20***
	61.40***	123.87***	126.39***
	57.70***	83.70***	76.24***
Constant	-1.445	-1.046	-0.869
	(1.031)	(1.365)	(1.393)
Observations	6,882	6,886	6,886

Note:
**** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1*
Robust standard errors (clustered on CCS category) in parentheses.
e(Collab – Mean^{IV}) and e(Collab – Mean$^{2\,IV}$) are the residuals from the first-stage regression which are introduced in the second-stage for the control function approach.

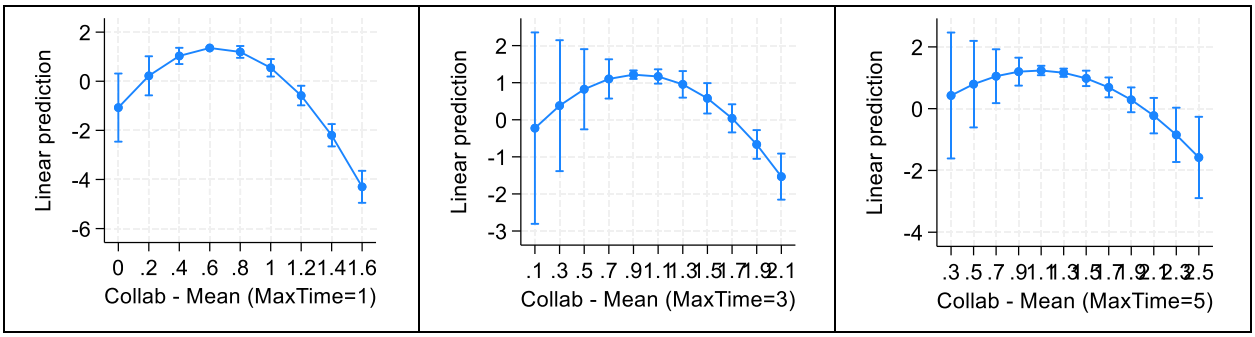

Figure C.4.2.: Marginal Effects for Non-surgical Services

C.5. Effect on 30-day Readmission

Table C.5.1.: Overall Collaborations and 30-day Readmission

	(1)	(2)	(3)
	$Max_{Time} = 1$	$Max_{Time} = 3$	$Max_{Time} = 5$
	Control Function (Probit)	**Control Function (Probit)**	**Control Function (Probit)**
Collab – Mean	13.771**	6.293	4.465
	(4.447)	(3.918)	(3.176)
(Collab – Mean)^2	-12.903***	-3.629*	-2.232*
	(3.586)	(1.792)	(1.168)
Collab – SD	0.67*	0.047	0.012
	(0.293)	(0.108)	(0.074)
e(Collab – Mean^{IV})	-13.319**	-5.251	-3.828
	(4.525)	(4.12)	(3.409)
e(Collab – Mean$^{2\,IV}$)	12.854***	3.435^	2.251^
	(3.51)	(1.828)	(1.243)
e(Collab – SD^{IV})	-0.682*	-0.061	-0.035
	(0.277)	(0.111)	(0.077)
Constant	-2.354*	-1.935	-1.583
	(0.945)	(1.574)	(1.833)
Observations	3,680	3,680	3,680

Note:
**** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1*
Robust standard errors (clustered on CCS category) in parentheses.
e(Collab – Mean^{IV}) and e(Collab – Mean$^{2\,IV}$) are the residuals from the first-stage regression which are introduced in the second-stage for the control function approach.

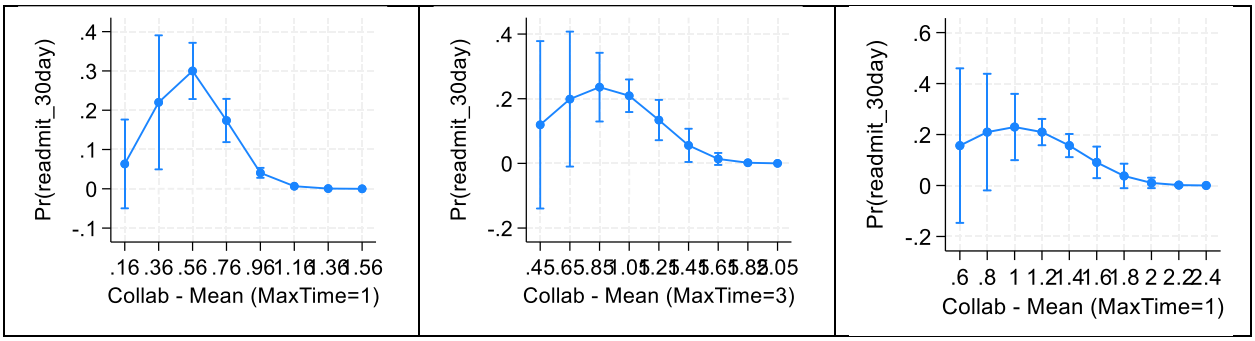

Figure C.5.2.: Predictive Marginal Effects for 30-day Readmission

C.6. Other Heterogeneity and Robustness Checks

Table C.6.1: Other Clustering Variables

	(1)	(2)	(3)	(4)
	$Max_{Time} = 1$	$Max_{Time} = 1$	$Max_{Time} = 1$	$Max_{Time} = 1$
	IV-2SLS	**IV-2SLS**	**IV-2SLS**	**IV-2SLS**
	(Week)	**(Week + CCS)**	**(Week + Adm.)**	**(ECI)**
Collab – Mean	8.653**	8.653**	8.653***	8.653*
	(2.786)	(3.13)	(2.046)	(4.416)
(Collab – Mean)^2	-6.999***	-6.999**	-6.999***	-6.999**
	(1.756)	(2.05)	(1.327)	(2.67)
Collab – SD	0.373**	0.373*	0.373**	0.373***
	(0.127)	(0.147)	(0.107)	(0.093)
***F*-test of Excluded Instruments**	46.37	34.97	29.92	79.04
	30.85	28.43	21.34	54.57
	17.99	20.42	19.81	24.04
Constant	-1.295^	-1.295	-1.295*	-1.295*
	(0.714)	(0.833)	(0.539)	(1.389)
Observations	8,402	8,402	8,402	8,402

Note:
**** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1*
Robust standard errors (clustered on CCS category) in parentheses.
e(Collab – Mean^{IV}) and e(Collab – Mean$^{2\,IV}$) are the residuals from the first-stage regression which are introduced in the second-stage for the control function (2SRI) approach.

Table C.6.2.: High and Low Patient Complexity

	(1)	(2)
	$Max_{Time} = 1$	$Max_{Time} = 1$
	IV-2SLS	**IV-2SLS**
	Low Complexity	**High Complexity**
Collab – Mean	3.382	37.32***
	(2.474)	(6.454)
(Collab – Mean)^2	-3.749*	-26.744***
	(1.484)	(4.542)
Collab – SD	0.521**	0.57**
	(0.157)	(0.205)
***F*-test of Excluded Instruments**	56.18	104.49
	45.51	72.88
	81.01	128.64
Constant	0.232	-10.564***
	(0.622)	(1.939)
Observations	6,091	2,311
Note: **** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1* *Robust standard errors (clustered on CCS category) in parentheses.*		

Table C.6.3.: Quantile Regressions for Robustness of Quadratic Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$Max_{Time} = 1$	$Max_{Time} = 1$	$Max_{Time} = 1$	$Max_{Time} = 3$	$Max_{Time} = 3$	$Max_{Time} = 3$	$Max_{Time} = 5$	$Max_{Time} = 5$	$Max_{Time} = 5$
	Control Function	Control Function	Control Function	Control Function	Control Function	Control Function	Control Function	Control Function	Control Function
	25th Quantile	50th Quantile	75th Quantile	25th Quantile	50th Quantile	75th Quantile	25th Quantile	50th Quantile	75th Quantile
Collab–Mean	8.352***	8.382***	6.297**	9.063***	4.121*	2.52	9.133***	2.667	0.854
	(2.147)	(1.948)	(2.382)	(1.68)	(1.642)	(1.971)	(1.992)	(1.865)	(2.168)
(Collab–Mean)^2	-6.448***	-6.202***	-5.479***	-3.743***	-1.96***	-1.535*	-3.117***	-1.166^	-0.674
	(1.367)	(1.24)	(1.516)	(0.669)	(0.654)	(0.77)	(0.671)	(0.628)	(0.73)

C.7. Post-Estimation Counterfactual Simulation

Table C.7.1: Post Estimation Counterfactual Results

Collaboration Strength (Max_{Time} = 1)	Counterfactual Increase in Collaboration	Change in LOS in Counterfactual
Lower Than Tipping Point (< 0.6)	0.1	15.71%
Higher Than Tipping Point (> 0.6)	0.1	-10.84%
All	0.1	0.35%
All	0.25	-2.32%
All	0.5 (1 S.D.)	-15.02%