Submitted to *Manufacturing & Service Operations Management*
manuscript MSOM-20-209.R2

Online Supplements - Service Chains' Operational Strategies: Standardization or Customization? Evidence from the Nursing Home Industry

Lu Kong
Muma College of Business, University of South Florida, Tampa, FL 33620, kongl@usf.edu

Kejia Hu
Owen Graduate School of Management, Vanderbilt University, Nashville, TN 37235, kejia.hu@vanderbilt.edu

Rohit Verma
VinUniversity, Hanoi, Vietnam & Cornell University, Ithaca, NY 14850, rohit.verma@cornell.edu

---

1. Control Variables

1. Local market-level characteristics: local market heterogeneity not only affects the three nursing home outcomes we explore, but also contributes to DoS decisions within a chain. To account for the influence of the local market, we add several location-level control variables and fixed effects:

(a) Average CM, SO, and SD for the local market each year. The base level of these variables may influence nursing home performance outcomes, independently of standardization strategies. Controlling these baseline variables helps disentangle location-specific variances from the general effects of deviation from the chain norm.

(b) The total number of nursing homes in the market each year. The number of local competitors can influence decisions about organization structures and performance outcomes (Castle et al. 2007).

(c) Market-level senior population, number of registered nurses, and average household income. The number of target customers, availability of labor, and local economic status of a location can influence the nursing home performance outcomes we explore.

(d) Location-level fixed effects. Adding location-level fixed effects soaks up systematic differences between markets that are time-invariant. We use Health Service Area (HSA) as a market/location level indicator in our analysis. The National Center for Health Statistics (part of the CDC) defines a health service area as one single county or cluster of contiguous counties that is relatively self-contained in terms of healthcare resources such as hospital care. The reasons we use HSA to define market level rather than other geographic regions, such as state, county, and zip code, are:

(d.1) The origination of HSA. The 805 HSAs are aggregated from counties by a cluster analysis of where the residents aged 65 years and over obtain their routine care. Since people frequently travel outside their counties and zip codes of residence for care, we argue that HSA as a geographic unit better captures the local market's population characteristics than state and county (especially in the context of our study, since the majority of nursing home customers are seniors).

(d.2) The measure of between-market variance. We argue that using the state as a local market indicator may mask interesting microlocal variations in the data, whereas county or zip code could exaggerate the between-market variances. The 805 HSAs are a well-defined compromise (in terms of size and number) between the 3,141 counties and 50 states.

2. Size: the number of facilities in a chain and the number of beds in a facility each year. The structure and outcomes of organizations are often contingent on size (Curry and Ratliff 1973).

3. Resident acuity level: resident average Activities of Daily Living (ADL) score and the average Resource Utilization Group Nursing Case Mix Index (RUGCMI) in each facility in each year. The average acuity level of residents affects the average level of care needed, and thus can be associated with multiple operational outcomes.

4. Staffing level: registered nurse hours per resident day in each facility each year. Staffing level is proven to be associated with clinical outcomes (Harrington et al. 2016) and nursing home costs.

5. Costs and occupancy: total costs and occupancy of each facility each year. Costs and occupancy can be associated with financial performance and resident welfare.

6. Property type: whether the focal facility is profit-based or hospital-based. Property type is found to be associated with various nursing home outcomes (Comondore et al. 2009).

7. Year fixed effects are added to capture unobservable effects that are time-variant.

8. Chain fixed effects are included in the model to account for time-invariant variance at the chain level, which may influence performance outcomes.

2. Table: Descriptive Analysis of Chains

Table 1 Descriptive Analysis of Chains

Chain Attributes	Criteria	Numbers
Chain Size by Facilities	The Largest Chain	104
	The Smallest Chain	3
	Average Chain Size	10
	Number of Large Chains (>= 20 Facilities)	15
	Number of Small Chains (<= 3 Facilities)	39
Chain Size by Total Beds	The Largest Chain	10,748
	The Smallest Chain	123
	Average Chain Size	1,124.2
	Number of Large Chains (>= 2000 Beds)	23
	Number of Small Chains (<= 300 Beds)	18
Chain Geographic Coverage	Number of Chains Operating in Only One State	74
	Number of Chains Operating in Multiple States	118
Chain Types	Number of Chains with Only For-profit Facilities	154
	Number of Chains with Only Non-profit Facilities	17
	Number of Chains with Both For-profit and Non-profit Facilities	21

3. Table: Correlations of Variables Used in Models

Table 2 Correlations of Variables Used in Models

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. *Operating Margin*	1.000																					
2. *Clinical Outcome*	-0.112	1.000																				
3. *Resident Welfare*	-0.059	0.379	1.000																			
4. *DoS in Service Offering*	-0.001	0.034	-0.004	1.000																		
5. *DoS in Service Delivery*	0.008	-0.034	0.021	-0.006	1.000																	
6. *DoS in Customer Mix*	0.150	-0.041	0.021	-0.005	0.029	1.000																
7. *Number of Facilities in Chain*	-0.038	-0.002	-0.022	-0.078	0.019	0.047	1.000															
8. *Average ADL*	0.081	-0.077	-0.023	0.022	0.099	-0.068	-0.022	1.000														
9. *Average RUGCMI*	-0.061	0.010	-0.049	0.012	0.133	-0.069	0.090	0.406	1.000													
10. *Registered Nurse Hours*	-0.078	-0.009	-0.015	0.003	-0.017	-0.158	0.123	0.196	0.406	1.000												
11. *Number of Beds in Facility*	0.067	-0.243	0.121	-0.114	0.065	0.078	-0.040	0.130	0.054	-0.001	1.000											
12. *Market Average Service Offering*	-0.022	-0.057	0.030	-0.301	0.020	0.006	0.097	-0.127	-0.102	0.026	0.051	1.000										
13. *Market Average Service Delivery*	-0.075	-0.012	0.035	0.017	0.130	-0.122	0.002	0.310	0.432	0.497	0.070	-0.038	1.000									
14. *Market Average Customer Mix*	0.011	-0.008	-0.012	-0.014	-0.019	0.242	0.013	-0.067	-0.044	-0.194	0.052	-0.056	-0.260	1.000								
15. *Cost*	-0.411	0.022	-0.026	0.011	-0.008	-0.119	-0.024	0.001	0.024	0.073	-0.073	0.021	0.046	-0.045	1.000							
16. *Occupancy*	0.204	-0.210	-0.036	0.007	0.092	0.018	0.093	0.208	0.001	0.053	-0.014	0.019	0.107	-0.073	0.012	1.000						
17. *Number of Facilities in Market*	-0.043	-0.065	0.034	0.025	0.028	-0.100	-0.051	0.108	0.064	0.188	0.240	0.040	0.150	-0.115	0.028	0.019	1.000					
18. *For-profit*	0.158	-0.033	0.030	-0.008	0.019	0.068	0.030	0.110	0.101	0.010	0.168	-0.081	0.070	0.177	-0.146	-0.030	0.042	1.000				
19. *Hospital-based*	0.003	-0.004	0.007	-0.009	0.009	-0.005	0.017	-0.001	0.020	0.035	0.001	0.014	0.021	0.000	-0.002	0.000	-0.015	0.008	1.000			
20. *Local Income*	-0.022	-0.035	0.080	0.005	0.055	-0.097	0.039	0.158	0.158	0.302	0.178	0.015	0.272	-0.229	0.024	0.100	0.397	0.024	0.005	1.000		
21. *Local Senior Population*	-0.008	-0.016	0.081	0.042	0.031	-0.102	-0.103	0.220	0.123	0.137	0.195	-0.051	0.212	-0.129	0.008	0.041	0.510	0.089	-0.006	0.224	1.000	
22. *Local Number of Registered Nurses*	0.012	0.017	-0.022	0.018	-0.021	-0.038	-0.133	0.185	0.116	-0.061	0.136	-0.182	0.075	0.065	-0.009	-0.065	0.213	0.151	-0.014	0.107	0.403	1.000

4. Tests and Validation for Choice of Instrumental Variables

We adopt an instrumental variable (IV) two-stage least squares (2SLS) approach, using lagged endogenous variables as IVs in our models. Despite its popularity, this approach is subject to some criticism. For example, Bellemare et al. (2017) find that lagging independent variables may lead to rejections of null hypotheses that are true; and this approach oftentimes (except when endogeneity threats and persistence in the unobservables are both small) generates estimations that are no better than simply ignoring the endogeneity. To guide proper use of this analysis approach, Bellemare et al. (2017) specify several conditions under which lagged explanatory variables are appropriate IVs in addressing endogeneity concerns. In this section, we validate to the extent possible that our analysis meets the conditions specified, and provide tests and results to support our choice of IVs.

4.1. The Relevance and Exclusion Conditions of IVs

In the nursing home industry, CMS's annual inspection provides valuable feedback to each nursing home, which is later used to adjust operating strategies for the following year. Thus, the yearly lagged explanatory variables should be correlated with the explanatory variables in the current year. This satisfies the "relevance" requirement of the IV approach. We test these hypothesized correlations with our first stage model, and the results (shown in Table 3) support our hypotheses.

Also, we assume the lagged explanatory variables are uncorrelated with the error terms, since the operational decisions in year $(t-1)$ should not determine the unobserved factors of operating outcomes in year $(t+1)$. This satisfies the "exclusion restriction" requirement of the IV approach. With evidence from simulations, Reed (2015) suggests that using lagged values of the endogenous variables as IVs is valid when (1) the IVs do not themselves belong in the respective estimating equation, and (2) the IVs are significantly correlated with the endogenous variables. Both of these requirements are met in our analysis.

4.2. Using Lagged Explanatory Variables as IVs

In their paper, Bellemare et al. (2017) specify several conditions; when satisfied, each of these conditions can independently justify the use of lagged explanatory variables as IVs. One of these conditions (condition 2a in the paper) has several requirements: (1) there is no unobserved confounding; (2) reverse causality does not exist ($Y \nrightarrow X$); (3) the causal effect only operates through the lagged explanatory variables ($X_{t-1} \rightarrow Y$ but $X_t \nrightarrow Y$). Since we add to our model a considerable amount of relevant control variables at the facility, chain, and market level, along with the chain, market, and year fixed effects, we argue that there is no unobserved confounding in our analysis. We then test for requirements (2) and (3). As stated in Section 4.2 of the main text, we

use outcome variables in year $(t+1)$ and explanatory variables (DoS in CM, SO, and SD) in year (t) before we justify our choice of IVs.

We first test for reverse causality. We regress each endogenous explanatory variable in year (t) on all outcome variables in our analysis, including operating margin, clinical outcome, and resident welfare. The control variables and fixed effects in equation (10) are also included in the reverse causality models. The results are shown in Table 4. As none of the outcome variables is significantly associated with any of the explanatory variables, we conclude that reverse causality does not exist in our analysis.

We then test if the causal effect only operates through the lagged explanatory variables (X_{t-1}), but not through the explanatory variables in year (t) (X_t). We regress the outcome variables on X_t and X_{t-1}, respectively; the control variables and fixed effects in equation (10) are also included in these models. The results are shown in Table 5 and Table 6. As indicated in Table 5, except for DoS in customer mix (significantly associated with operating margin), no other explanatory variable in year (t) is significantly associated with any outcome variables. Hence, out of nine hypothesized relationships (three explanatory variables times three outcome variables), eight satisfy the condition specified in the Bellemare et al. (2017) paper. Meanwhile, Table 6 exhibits several significant relationships between the lagged explanatory variables (X_{t-1}) and the outcome variables. Thus, we argue that in our analysis, the condition specified by Bellemare et al. (2017) is met to a sufficient extent to justify using lagged explanatory variables as IVs.

4.3. Models without IVs

To help readers better understand the direction and magnitude of the biases and the impact of the instrumental variable analysis itself, we present the results of models without instrumental variables in Table 7. Results indicate that several important findings are masked in models without instrumental variables, such as the significant relationship between DoS in customer mix and clinical outcome, and the marginally significant relationship between DoS in service delivery and resident welfare. Also worth noticing is that the IV models lose observations from lagging variables, but they also have higher adjusted R^2 than the models without IVs.

Table 3 First Stage Model Results 1

	First Stage Model 1 Service Offering	First Stage Model 2 Service Delivery	First Stage Model 3 Customer Mix
Constant	-1880.0000	-6290.0000***	-1700.0000***
	(1400.0000)	(299.0000)	(168.0000)
Explanatory Variables			
Lagged Service Offering	5680.0000***	-	-
	(63.0000)	-	-
Lagged Service Delivery	-	4250.0000***	- .
	-	(63.8000)	-
Lagged Customer Mix	-	-	7400.0000***
	-	-	(47.0000)
Facility-Level Control Variables			
Average ADL	-103.0000·	16.2000	-0.7150
	(55.2000)	(11.5000)	(6.5700)
Average RUGCMI	1150.0000	1360.0000***	-234.0000
	(1370.0000)	(287.0000)	(163.0000)
Registered Nurse Hours	-399.0000	-1370.0000***	-278.0000***
	(464.0000)	(97.0000)	(55.3000)
Number of Beds in Facility	-17.2000***	1.2100*	0.8970**
	(2.3900)	(0.4970)	(0.2840)
Occupancy	0.8300	10.5000***	3.5200***
	(7.4400)	(1.5600)	(0.8860)
For-profit	-33.3000	50.6000	129.0000**
	(343.0000)	(71.6000)	(40.8000)
Hospital-based	-232.0000	751.0000	97.3000
	(4790.0000)	(1000.0000)	(570.0000)
Cost	0.0032	-0.0008	-0.0045***
	(0.0061)	(0.0013)	(0.0007)
Chain-Level Control Variables			
Number of Facilities in Chain	-4.6100**	0.4110	0.3840*
	(1.4200)	(0.2950)	(0.1680)
Location/Market-Level Control Variables			
Market Average Service Offering	-12700.0000***	153.0000.	77.3000
	(453.0000)	(91.8000)	(52.3000)
Market Average Service Delivery	561.0000	878.0000***	-7.6400
	(421.0000)	(88.1000)	(50.1000)
Market Average Customer Mix	-2.3200	1.8600	14.1000***
	(10.2000)	(2.1200)	(1.2300)
Number of Facilities in Market	51.9000***	4.7600.	-2.0200
	(13.7000)	(2.8500)	(1.6300)
Local Income	0.0072	0.0051**	0.0015
	(0.0088)	(0.0018)	(0.0011)
Local Senior Population	0.0014	0.0002	-0.0003**
	(0.0009)	(0.0002)	(0.0001)
Local Number of Registered Nurses	-0.0079***	-0.0024***	-0.0002
	(0.0019)	(0.0004)	(0.0002)
Year Fixed Effect	Yes	Yes	Yes
R^2	0.3842	0.2391	0.6188
Adjusted R^2	0.3834	0.2380	0.6183
Degree of Freedom	18322	18312	18322
F-statistic	457.3***	230.1***	1190***

Notes. The reported coefficients and standard errors were multiplied by 10^4

$.p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 4 Reverse Causality Models ($Y \rightarrow X_t$)

	Model 1 Service Offering	Model 2 Service Delivery	Model 3 Customer Mix
Operating Margin	Insignificant (p= 0.35)	Insignificant (p= 0.29)	Insignificant (p= 0.21)
Clinical Outcome	Insignificant (p= 0.35)	Insignificant (p= 0.21)	Insignificant (p= 0.85)
Resident Welfare	Insignificant (p= 0.17)	Insignificant (p= 0.88)	Insignificant (p= 0.61)
Facility-Level Control Variables	Yes	Yes	Yes
Chain-Level Control Variables	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes
Chain-Level Fixed Effects	Yes	Yes	Yes
Location/Market-Level Fixed Effects	Yes	Yes	Yes
R^2	0.2748	0.2241	0.4959

Table 5 Explanatory Variables and Outcome Variables ($X_t \rightarrow Y$)

	Model 1 Operating Margin	Model 2 Clinical Outcome	Model 3 Resident Welfare
DoS in Service Offering	Insignificant (p= 0.27)	Insignificant (p= 0.81)	Insignificant (p= 0.19)
DoS in Service Delivery	Insignificant (p= 0.10)	Insignificant (p= 0.19)	Insignificant (p= 0.78)
DoS in Customer Mix	***Significant*** (p= 0.00)	Insignificant (p= 0.76)	Insignificant (p= 0.82)
Facility-Level Control Variables	Yes	Yes	Yes
Chain-Level Control Variables	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes
Chain-Level Fixed Effects	Yes	Yes	Yes
Location/Market-Level Fixed Effects	Yes	Yes	Yes
R^2	0.3793	0.2562	0.2952

Table 6 One-Period Lagged Explanatory Variables and Outcome Variables ($X_{t-1} \rightarrow Y$)

	Model 1 Operating Margin	Model 2 Clinical Outcome	Model 3 Resident Welfare
Lagged DoS in Service Offering	Insignificant (p= 0.43)	Insignificant (p= 0.70)	Insignificant (p= 0.42)
Lagged DoS in Service Delivery	***Significant*** (p= 0.03)	Insignificant (p= 0.63)	***Significant*** (p= 0.06)
Lagged DoS in Customer Mix	***Significant*** (p= 0.00)	***Significant*** (p= 0.02)	Insignificant (p= 0.70)
Facility-Level Control Variables	Yes	Yes	Yes
Chain-Level Control Variables	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes
Chain-Level Fixed Effects	Yes	Yes	Yes
Location/Market-Level Fixed Effects	Yes	Yes	Yes
R^2	0.4017	0.2753	0.3038

Table 7 Models without Instrumental Variables

	Model 1 Operating Margin	Model 2 Clinical Outcome	Model 3 Resident Welfare
Constant	-2474.0000***	351.4000***	61440.0000**
	(681.2000)	(91.3800)	(22590.0000)
Explanatory Variables			
DoS in Service Offering	-6.9630	-0.2075	272.6000
	(6.3190)	(0.8476)	(209.5000)
DoS in Service Delivery	-50.0000.	-5.3100	-279.7000
	(30.2900)	(4.0640)	(1004.0000)
DoS in Customer Mix	172.6000***	-1.9710	360.0000
	(47.1900)	(6.3300)	(1565.0000)
Facility-Level Control Variables			
Average ADL	26.6400***	-0.2825	202.0000
	(6.0310)	(0.8094)	(200.0000)
Average RUGCMI	375.6000*	17.4300	-4662.0000
	(152.2000)	(20.4300)	(5048.0000)
Registered Nurse Hours	-68.2100	-19.1600*	-10580.0000***
	(56.4100)	(7.5690)	(1870.0000)
Number of Beds in Facility	2.2120***	-0.7838***	149.3000***
	(0.2874)	(0.0386)	(9.5310)
Occupancy	20.6200***	-2.0250***	-128.5000***
	(0.9668)	(0.1297)	(32.0600)
For-profit	884.6000***	15.2500	3959.0000
	(91.6200)	(12.3200)	(3038.0000)
Hospital-based	970.8000*	-124.6000*	-46850.0000**
	(438.2000)	(58.7900)	(14530.0000)
Cost	-0.0210***	0.0000	-0.0245
	(0.0008)	(0.0001)	(0.0272)
Chain-Level Control Variables			
Number of Facilities in Chain	-3.0730***	-0.0397	70.6200***
	(0.5726)	(0.0768)	(18.9900)
Location/Market-Level Control Variables			
Market Average Service Offering	-164.4000*	-2.8010	125.7000
	(72.5800)	(9.7370)	(2407.0000)
Market Average Service Delivery	40.4300	-14.4200*	-341.9000
	(54.7600)	(7.3470)	(1816.0000)
Market Average Customer Mix	-1.6380	-0.4282.	-52.3900
	(1.8140)	(0.2434)	(60.1600)
Number of Facilities in Market	-7.6690	0.1014	19.0700
	(5.1100)	(0.6856)	(169.4000)
Local Income	0.0005	0.0000	0.0342
	(0.0015)	(0.0002)	(0.0507)
Local Senior Population	-0.0005*	0.0000	0.0233***
	(0.0002)	(0.0000)	(0.0070)
Local Number of Registered Nurses	-0.0032	0.0002	0.0842
	(0.0023)	(0.0003)	(0.0753)
Fixed Effects			
Year Fixed Effect	Yes	Yes	Yes
Chain-Level Fixed Effects	Yes	Yes	Yes
Location/Market-Level Fixed Effects	Yes	Yes	Yes
R^2	0.3793	0.2562	0.2952
Adjusted R^2	0.3466	0.2171	0.2581
Degree of Freedom	17387	17381	17387
F-statistic	11.62***	6.551***	7.967***

Notes. The reported coefficients and standard errors were multiplied by 10^4

$.p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

References

Bellemare MF, Masaki T, Pepinsky TB (2017) Lagged explanatory variables and the estimation of causal effect. *The Journal of Politics* 79(3):949–963.

Castle NG, Engberg J, Liu D (2007) Have nursing home compare quality measure scores changed over time in response to competition? *BMJ Quality & Safety* 16(3):185–191.

Comondore VR, Devereaux P, Zhou Q, Stone SB, Busse JW, Ravindran NC, Burns KE, Haines T, Stringer B, Cook DJ, et al. (2009) Quality of care in for-profit and not-for-profit nursing homes: systematic review and meta-analysis. *BMJ* 339(b2732).

Curry TJ, Ratliff BW (1973) The effects of nursing home size on resident isolation and life satisfaction. *The Gerontologist* 13(3_Part_1):295–298.

Harrington C, Schnelle JF, McGregor M, Simmons SF (2016) Article commentary: The need for higher minimum staffing standards in us nursing homes. *Health Services Insights* 9:13–19.

Reed WR (2015) On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics* 77(6):897–905.