Journal of Behavioral
Public Administration
Vol 3(1), pp. 1-8
DOI: 10.30636/jbpa.31.107

Research Article

Sector bias in public programs: US nonprofit hospitals

Kenneth J. Meier*, Seung-Ho An†

Abstract: This study investigates whether the public perceives nonprofit organizations as different from private for-profit and public organizations and whether introducing new performance management systems would provide positive credits to the organization. Using two randomized survey experiments on US hospitals (one with an adult sample and the other with a student sample), we replicate the study of Hvidman & Andersen (2016) in Denmark with an extension of adding a nonprofit organization cue. The results show no sectoral differences among the hospitals and no positive feedback for adopting a new performance management system.

Keywords: Sectoral differences, Sector bias, Performance management, Replication, Nonprofit organization

Supplements: Open data, Open materials

The public's ability to evaluate government programs is an essential part of the feedback process that links individual citizens to elected officials in the democratic process. Of major concern, however, is that the perceptual biases of individuals that are well documented in the behavioral public administration literature preclude objective evaluation (Belle, Cantarelli, & Belardinelli, 2018; James, Jilke, Petersen, & Walle, 2016; Marvel, 2015; Olsen, 2017). Among the findings in the literature are that there is a perceived bias against public organizations, that is, public organizations are evaluated more negatively than private organizations even when the objective outcomes are identical (Marvel, 2015; Hvidman & Andersen, 2016). Such findings are not without challenge. Meier, Johnson, & An (2019), replicating a Danish study (Hvidman & Andersen 2016) of hospitals that found bias, showed no sector bias in two replications in the US. Experimental subjects in the replication study rated US public hospitals no different from private hospitals.

Differences in findings of the two studies can result from differences in context, differences in subjects, or differences in the structure of the experiments (Walker, James & Brewer, 2017). Given the similar design of both experiments and given the US study used two different subject pools (one MTurk adults and one students; the Danish study used students), differences in context are the likely explanation. Among the differences in delivery of public programs via hospitals in Denmark and the US is that the US hospital sector is dominated by nonprofit hospitals (62%); public hospitals total 20% and for-profit hospitals 18%. In Denmark, 97% of hospital beds are publicly owned. To keep the replication of Hvidman and Andersen as parallel as possible, Meier et al. (2019) only reported on the public versus private differences in the US setting.

Studies examining sector bias in public programs need to include nonprofit organizations for both empirical and theoretical reasons. In the United States as well as in many developing countries, public programs are often implemented by nonprofit organizations rather than by government agencies or private sector organizations. In addition to hospital care, nonprofits play a major role in implementing

* American University, Cardiff University, Leiden University
† University of Arizona
Address correspondence to Kenneth J. Meier at (kmeier@american.edu)

programs in education (charter schools and private universities), substance abuse treatment, mental health services, housing, and family planning, among others. Omitting such organizations could lead to misleading results. Theoretically, nonprofit organizations rely heavily on normative incentives and altruistic motivation (Donahue & Zechhauser 2012), and this orientation might create positive halo effects that favor nonprofit programs over those delivered by government agencies. This might be especially the case in assessing performance criteria such as benevolence.

This study contributes to the literature in terms of empirical generalization and extension by replicating Hvidman & Andersen (2016); in other words, we report a nonprofit organization cue that was included in the Meier et al. (2019) study. Our results, which are based on all three sectors (public, private, and nonprofit), reveal no differences in evaluations based on the sector of the hospital. Given that the results also show no differences across the three sectors, the difference in results between the US and Denmark studies cannot be attributed to the sector differences in implementing health policy and is likely to come from other contextual factors that differentiate the two countries.

The Experiment

Hvidman & Andersen (2016) presented a basic 2×2 factorial design using university undergraduates. The subjects were exposed to vignettes about a hospital that was designated as either public or private; a second experimental manipulation randomly provided half the respondents with a performance evaluation cue, that is, the hospital contracted with an outside consultant to evaluate the hospital and this evaluation produced positive results. Our US replication was a 3×2 design with respondents randomly assigned to the hospital being public, private, or nonprofit and then randomly assigned to the same performance evaluation cue.[1] Two replications were co-

Table 1
Balance across Experimental Groups

Adult Sample		Ideology	Age	Religious Service Attendance	Gender	White
Public	Baseline	2.50	37.38	1.95	0.48	0.75
	PM	2.78	35.97	2.01	0.53	0.67
Private	Baseline	2.46	35.15	1.88	0.53	0.77
	PM	2.38	35.87	1.89	0.52	0.77
Nonprofit	Baseline	2.58	36.55	1.92	0.62	0.80
	PM	2.68	36.13	1.97	0.51	0.75
	Prob. > F	0.19	0.91	0.97	0.59	0.46
Student Sample		Ideology	Age	Religious Service Attendance	Gender	White
Public	Baseline	3.37	19.05	1.95	0.45	0.58
	PM	3.10	18.92	1.90	0.35	0.59
Private	Baseline	3.38	19.18	1.98	0.41	0.72
	PM	3.28	19.04	1.85	0.52	0.73
Nonprofit	Baseline	3.30	18.82	1.97	0.44	0.71
	PM	3.12	18.94	1.99	0.43	0.63
	Prob. > F	0.14	0.24	0.93	0.32	0.06

Note: PM indicates the performance management cue group; ideology is on a five-point scale from 1 (very liberal) to 5 (very conservative); religious service attendance is on a four-point scale from 1 (never) to 4 (weekly); for gender and White, 1 indicates female and white, respectively (otherwise, zero); age is a continuous variable.

Table 2
Factor Analysis of the Four Perceived Performance Dimensions

Survey Item (Adult Sample)	Effectiveness	Red Tape	Efficiency	Benevolence
The hospital is effective.	**0.78**	0.03	0.02	0.26
The hospital is effective in accomplishing its core mission.	**0.84**	0.00	0.03	0.23
The hospital is effective in delivering a very good service.	**0.80**	0.02	0.05	0.25
The hospital has disproportionate administrative rules and procedures.	-0.01	**0.66**	-0.02	0.13
A high level of administrative procedures negatively affects the hospital's efficiency.	0.06	**0.68**	-0.04	0.19
The hospital is effective in lowering its costs.	0.30	-0.06	**0.33**	0.10
The hospital prioritizes its profit margin.	0.12	-0.25	**0.32**	-0.16
The hospital is genuinely interested in the well-being of patients.	0.24	0.12	0.00	**0.85**
The hospital acts in the interest of patients.	0.30	0.09	-0.01	**0.85**
Cronbach's alpha	0.90	0.73	0.35	0.92
Mean	0	0	0	0
Standard deviation	0.91	0.76	0.43	0.90
N	554	554	554	554
Survey Item (Student Sample)	Effectiveness	Red Tape	Efficiency	Benevolence
The hospital is effective	**0.69**	-0.14	0.00	0.25
The hospital is effective in accomplishing its core mission.	**0.72**	-0.07	-0.04	0.25
The hospital is effective in delivering a very good service.	**0.68**	-0.07	0.00	0.28
The hospital has disproportionate administrative rules and procedures.	-0.14	**0.53**	0.12	-0.16
A high level of administrative procedures negatively affects the hospital's efficiency.	-0.22	**0.53**	0.06	-0.15
The hospital is effective in lowering its costs.	-0.01	0.09	**0.43**	-0.04
The hospital prioritizes its profit margin.	-0.03	0.20	**0.43**	-0.12
The hospital is genuinely interested in the well-being of patients.	0.24	-0.06	-0.04	**0.78**
The hospital acts in the interest of patients.	0.25	-0.10	-0.03	**0.78**
Cronbach's alpha	0.82	0.62	0.46	0.86
Mean	0	0	0	0
Standard deviation	0.83	0.64	0.54	0.84
N	638	638	638	638

Table 3
Perceptions of Performance:
Impact of Sector and Performance Management, Regression Models (OLS)

Adult Sample	Effectiveness	Red Tape	Efficiency	Benevolence
Private Organizations (Private=1)	-0.051	0.043	-0.024	-0.016
	(0.095)	(0.078)	(0.047)	(0.097)
Public Organizations (Public=1)	-0.058	0.104	-0.055	0.082
	(0.096)	(0.078)	(0.045)	(0.095)
Performance Management	-0.008	0.014	-0.007	0.229*
	(0.077)	(0.065)	(0.036)	(0.076)
Constant	0.042	-0.058	0.031	-0.137
	(0.078)	(0.062)	(0.040)	(0.080)
R-Squared Overall	0.0008	0.0033	0.0029	0.0187
N	554	554	554	554
Student Sample	Effectiveness	Red Tape	Efficiency	Benevolence
Private Organizations (Private=1)	0.033	0.000	0.046	-0.095
	(0.078)	(0.061)	(0.051)	(0.082)
Public Organizations (Public=1)	0.004	-0.008	0.014	-0.113
	(0.084)	(0.062)	(0.053)	(0.081)
Performance Management	-0.101	0.002	-0.046	-0.053
	(0.066)	(0.051)	(0.042)	(0.067)
Constant	0.037	0.001	0.002	0.093
	(0.068)	(0.050)	(0.041)	(0.068)
R-Squared Overall	0.0041	0.0000	0.0033	0.0042
N	638	638	638	638

Note: * $p<0.05$; robust standard errors in parentheses; two-tailed tests of significance; nonprofit organizations are the reference group across models.

nducted.[2] The first replication used 554 adult subjects selected from Amazon MTurk; the second replication used 638 undergraduate students from a large public university.[3] The full documentation of the experiment can be found in the Appendix.

Findings

Tables 1 to 4 include the results from the two randomized survey experiments using both adult and student samples; each table presents the results of adult and student sample above and below, respectively.

Table 1 reports the balance tests for the two experiments comparing the experimental and control groups on ideology, age, religious service attendance, gender, and race. None of the F-tests for imbalance are statistically significant, indicating that the randomization was effective. Table 2 shows the factor analyses for the dependent variables – effectiveness, red tape, efficiency, and benevolence for each of the sample pools. The factor analyses are slightly different from those appearing in Meier et al. (2019) owing to the larger sample size that results from including the subjects with the nonprofit treatment. The new samples, however, produce factor results that are very close to the original set, including the problematic 0.35 Cronbach's alpha for the efficiency measure. Both the balance test and the factor analysis results indicate that the nonprofit group should be comparable to the public sector and private sector set of respondents.

Table 4
Perceptions of Performance:
Moderating Impact of Performance Management, Regression Models (OLS)

Adult Sample	Effectiveness	Red Tape	Efficiency	Benevolence
Private Organizations (Private=1)	-0.105	0.023	-0.078	-0.013
	(0.126)	(0.108)	(0.061)	(0.136)
Public Organizations (Public=1)	0.052	0.050	-0.050	-0.010
	(0.129)	(0.110)	(0.066)	(0.134)
Performance Management (PM)	0.035	-0.041	-0.042	0.161
	(0.141)	(0.108)	(0.069)	(0.145)
Private × PM	0.102	0.044	0.109	0.001
	(0.190)	(0.156)	(0.094)	(0.195)
Public × PM	-0.218	0.111	-0.005	0.187
	(0.193)	(0.157)	(0.090)	(0.190)
Constant	0.022	-0.033	0.047	-0.106
	(0.095)	(0.074)	(0.051)	(0.098)
R-Squared Overall	0.0065	0.0042	0.0067	0.0212
N	554	554	554	554
Student Sample	Effectiveness	Red Tape	Efficiency	Benevolence
Private Organizations (Private=1)	0.058	-0.064	0.022	-0.125
	(0.110)	(0.085)	(0.069)	(0.116)
Public Organizations (Public=1)	0.094	-0.035	0.016	-0.106
	(0.117)	(0.087)	(0.065)	(0.110)
Performance Management (PM)	-0.031	-0.057	-0.060	-0.068
	(0.113)	(0.085)	(0.074)	(0.117)
Private × PM	-0.043	0.125	0.047	0.060
	(0.156)	(0.123)	(0.102)	(0.164)
Public × PM	-0.185	0.048	-0.009	-0.019
	(0.168)	(0.123)	(0.106)	(0.164)
Constant	-0.001	0.033	0.010	0.102
	(0.085)	(0.061)	(0.047)	(0.086)
R-Squared Overall	0.0063	0.0018	0.0038	0.0046
N	638	638	638	638

Note: * $p<0.05$; robust standard errors in parentheses; two-tailed tests of significance; nonprofit organizations are the reference group across models.

In the original US experiment, the basic evaluation of public hospitals was not statistically different from the respondents' evaluations of private hospitals. Findings from Table 3 using the adult and student sample show consistent results. There are no statistically significant differences in the evaluation of the hospitals based on sector; nonprofit hospitals are not rated either better or worse than public and

private hospitals in terms of effectiveness, red tape, efficiency, or benevolence.

Both experiments followed the lead of Hvidman & Andersen (2016) using an experimental manipulation of the hospitals' use of performance information; the vignette stated that the hospital hired an outside consultant to assess the hospital's performance and that assessment reported positive results (for more details, see Appendix). The hypothesis was that public hospitals would not get credit for good performance, but private hospitals would. The performance manipulation had no impact on the evaluation of the Danish hospitals whether public or private, and the US study showed similar insignificant results with the exception of a positive change in terms of benevolence in the adult sample (but not the student sample) before the experimental condition was assessed by sector.

Table 3 reports the impact of the performance appraisal cue controlling for the sector of the hospital, and Table 4 interacts the performance appraisal cue by sector. With the exception of the benevolence impact (direct impact only) for the adult population (Table 3), none of the results are statistically different from zero; the direct impact of performance management on benevolence even disappears when introducing interaction terms (Table 4). The basic conclusion is that US hospitals get no credit for engaging in an external performance appraisal that shows positive results, and this lack of results is consistent across the public, private, and nonprofit sectors. Overall the experiments found no sector biases in the US hospital sector, either as a direct effect or in interaction with a performance appraisal. Nonprofit hospitals were not different from either public or private hospitals.

Conclusion

Our replication of a Danish sector bias study involving hospitals failed to show any sector biases; public, private, and nonprofit hospitals with similar performance were rated no different from each other in two samples of subjects. Given that generally the perception is that the US has a pro-private sector orientation, additional research is clearly merited. One possibility for the absence of perceived differences, in fact, might be that at least one observational study of US hospitals found that on average they did not differ in terms of a wide range of performance indicators that included both the criteria for good medical treatment, hospital readmission rates, mortality rates, and public evaluations of the hospitals (Cheon, Song, McCrea & Meier, 2019). The similar performance ratings across US hospitals might have diluted the sectoral biases of the public since the public organizations are not necessarily performing poorly compared to other types of organizations. In other words, the similar performance ratings of hospitals across sectors in the US may contribute to our null findings. Other possibilities that might reveal biases should be considered such as the lack of negative information (the only cue was the positive performance management cue), the credibility of the performance evaluation information (the current cues were not especially precise), or the inclusion of other evaluation criteria that are more salient to the general public. We encourage future scholars to examine the above mentioned possibilities, replicating and extending the sector bias experiment.

Notes

1. To bolster the effects of sector cue, we included an addition, who controls the hospital, to the original sector cue by Hvidman & Andersen (2016); for more details on the sector cues, see Appendix. Nonetheless, our findings show no significant results.
2. The replication files are available at Havard Dataverse: https://doi.org/10.7910/DVN/KFIZJZ.
3. About 43 percent students failed to pass our manipulation check (identifying the sector), which is 19 percent higher than Hvidman & Andersen (2016). Excluding the subjects who failed to correctly identify the sector cue, the results are largely the same. Since excluding subjects that failed to pass the manipulation check can potentially lead to bias in estimation (for more details see Aronow, Baron, & Pinson, 2019), we report the results with full sample.

References

Aronow, P. M., Baron, J. & Pinson, L. (2019). A note on dropping experimental subjects who fail a manipulation check. *Political Analysis, 27*(4), 572-589.

Bellé, N., Cantarelli, P., & Belardinelli, P. (2018). Prospect theory goes public: Experimental evidence on cognitive biases in public policy and management decisions. *Public Administration Review, 78*(6), 828-840.

Cheon, O., Song, M., McCrea, A. M., & Meier, K. J. (2019). Health care in America: The relationship between subjective and objective assessments of hospitals. *International Public Management Journal*, doi/full/10.1080/10967494.2018.1553808.

Donahue, J.D. & Zeckhauser, R.J. (2012). *Collaborative governance: Private roles for public goals in turbulent times*. Princeton, NJ: Princeton University Press.

Hvidman, U., & Andersen, S. C. (2016). Perceptions of public and private performance: Evidence from a survey experiment. *Public Administration Review, 76*(1), 111-120.

James, O., Jilke, S., Petersen, C., & Walle, S. V. d. (2016). Citizens' blame of politicians for public service failure: Experimental evidence about blame reduction through delegation and contracting. *Public Administration Review, 76*(1), 83-93.

Marvel, J. D. (2015). Unconscious bias in citizens' evaluations of public sector performance. *Journal of Public Administration Research and Theory, 26*(1), 143-158.

Meier, K. J., Johnson, A., & An, S. (2019). Perceptual bias and public programs: The case of the United States and hospital care. *Public Administration Review, 79*(6), *820-828*.

Olsen, A. L. (2017). Compared to what? Experimental evidence on social and historical reference points in performance evaluation. *Journal of Public Administration Research and Theory, 27*(4), 562-580.

Walker, R. M., James, O., & Brewer, G. A. (2017). Replication, experiments and knowledge in public management research. *Public Management Review, 19*(9): 1221-1234.

Appendix

Vignette

The exact wording of our vignette for both MTurk and student sample is as follows:

Atlas Hospital, [sector cue], is controlled by [sector cue] and has 450 employees. The hospital is organized into three sections. The hospital's top management function is performed by a management team consisting of a medical director, a chief nursing officer, and a hospital director. Atlas Hospital's central administrative department is responsible for documenting that the hospital meets management's demands for safe and effective treatment. This task involves a comprehensive system of policies and standards in all areas of the hospital. The hospital's goal is to provide health care of the highest quality, taking into account the special needs of the individual patients. [performance cue] The hospital has hired a consulting firm to give an overall assessment of the organization. The consultants concluded that:

"The hospital uses modern technological equipment for examination and treatment."

"The staff has a high level of professionalism and regularly participate in training activities around the world."

Performance cue

The hospital has introduced a performance management process in which performance information is collected and then returned to managers, who use it to adjust the objectives. This process involves the hospital continuously collecting data on the quality of treatment and patient satisfaction, data that is then used for evaluations of the hospital's effectiveness.

Sector cue
1) Public: A county hospital, Atlas County
2) Private: A private company, Atlas Health Services, Inc.
3) Nonprofit: A non-profit institution, Atlas Charities Group