# Sector Bias and the Credibility of Performance Information:

# An Experimental Study of Elder Care Provision

Kenneth J. Meier
American University, Cardiff University, Leiden University
Department of Public Administration and Policy
School of Public Affairs
American University
Washington DC 20016
kmeier@american.edu

Miyeon Song
Rutgers University – Newark
School of Public Affairs and Administration
111 Washington Street
Newark, NJ 07102
misong@sc.edu

Jourdan Davis
University of North Carolina at Charlotte
Department of Political Science and Public Administration
9201 University City Blvd
Charlotte, NC 28223
j.davis@uncc.edu

Anna Amirkhanyan
American University
Department of Public Administration and Policy
School of Public Affairs
American University
Washington DC 20016
amirkhan@american.edu

forthcoming Public Administration Review

Earlier versions of this paper were presented at the 2019 Association for Research on Nonprofit Organizations and Voluntary Action Fall Annual Meeting, November 21-23, 2019, San Diego CA and at the 2019 meetings of the Association for Public Policy Analysis and Management, November 7-9, Denver Colorado. We would like to thank Saar Alon-Barkat for helpful comments.

# Author Bios

**Kenneth J. Meier** is a Distinguished Scholar in Residence at the School of Public Affairs, American University, Professor of Bureaucracy and Democracy at Leiden University (the Netherlands), and a Professor of Public Management at the Cardiff School of Business, Cardiff University (Wales). His research interests include social equity, public management, the role of bureaucracy in democratic systems, comparative public administration, and behavioral approaches to public administration.

**Miyeon Song** is an Assistant Professor in the School of Public Affairs and Administration at Rutgers University - Newark. Her research interests include public management, organizational theory and behavior, government performance, citizen satisfaction, and comparative public administration.

**Jourdan Davis** is an Assistant Professor in the Department of Political Science and Public Administration at University of North Carolina at Charlotte. Her research interests include behavioral public administration, fairness, justice, citizen evaluations, and public policy.

**Anna Amirkhanyan** is a professor of public administration and policy in the School of Public Affairs, American University. Her research on public and nonprofit management, performance, privatization, and citizen participation has been published in various peer-reviewed outlets such as the *Journal of Public Administration Research and Theory*, *Public Administration Review*, and others. Her book "Citizen Participation in the Age of Contracting" examines the state of citizen participation in the context of health and human services.

## Abstract

Reporting government performance to the public is key tool in improving accountability. Some evidence, however, has shown that individuals' anti-public sector biases may distort performance information about public organizations. Using an experimental vignette on U.S. nursing homes, this study fills four gaps in the literature: 1) the need to include nonprofit organizations rather than just public and for-profit, 2) consideration of the credibility of the source of performance information, 3) the use of simple commonly used performance metrics, and 4) the willingness to use services as a performance dimension. We find the public has a general but modest anti-for-profit sector bias in nursing home care with nonprofits perceived the most positively. Sector biases generally disappear when clear performance data are presented. The credibility of the source matters, and respondents' willingness to use organizational services is more sensitive to both sector bias and performance ratings than are performance measures.



***Keywords***: Sector bias, Anti-government bias, Performance information, Credibility


## Practitioner Points

- In the field of long-term care, individuals have no anti-public sector bias, and they have more favorable views of nonprofit organizations than for-profit organizations.
- Perceived sector bias disappears when simple and unambiguous performance information is introduced.
- The source of performance information matters: individuals perceive the information from government and nonprofit sources as more credible than that from for-profit sources.

- A high degree of credibility amplifies the effect of performance information in shaping individuals' evaluations of public services.

# Introduction

Public attitudes and perceptions of government and nonprofit organizations have crucial implications for the wellbeing of a nation's economy and society as a whole. In a democracy, citizens[1] are expected to engage closely with government by paying taxes, voting, commenting on new laws, expressing their feedback on existing services, and co-producing goods and services. At the same time, public sector performance has long suffered from negative public perceptions (Bok 2001; Roberts 2019). A Pew Research Center (2019) survey shows that only 17% of Americans trust the government in Washington D.C. to do what is right, a figure near historic lows. Since the empirical evidence has not supported the idea that private organizations outperform their public counterparts across several policy areas (Amirkhanyan, Kim, and Lambright 2008; Andrews, Boyne, and Walker 2011; Goodsell 2015; Hodge 2018), some scholars have asked whether the public holds biased perceptions of government organizations and gives them less credit for their work.

The growing literature on behavioral public administration has provided some initial evidence on individuals' inability or unwillingness to favorably evaluate public programs (Baekgaard and Sterritzlew 2016, 2018; Battaglio et al. 2018; Charbonneau and Van Ryzin 2015; Hvidman 2019; Hvidman and Andersen 2016; James and Van Ryzin 2017; Marvel 2015, 2016; Zhang et al. 2021). Some studies suggest that anti-public sector bias, involving perceptions of public organizations as wasteful and ineffective, creates cynicism that is resistant to objective information, especially among those with less education (Marvel, 2015, 2016).

While engaging with service organizations, citizens often rely on prior assessments of organizational performance. Whether they select a nonprofit to contribute to or make a purchase on Amazon, people are likely to check what others think about these experiences. In the 21st

Century, we are less likely to rely on the word of mouth for these opinions. Instead, customer reviews (the "virtual" word of mouth), numeric indicators, and quality ratings, which aggregate a large number of assessments from experts or prior users are commonplace (Olsen 2017a). Such performance information can come from nonprofit and for-profit sources, as well as from government entities charged with oversight and regulation. Evidence on whether individuals view performance information from government regulators as more or less credible than similar information coming from nonprofit professional groups or for-profit sources involving customer ratings is limited and inconclusive.

Public and private organizations collect much data on vitally important services, trends and outcomes; so the question of credibility of performance information has immense practical importance. While performance assessment is driven by numerous factors, one key objective is to inform individual decision-making (Moynihan 2006). Biases against certain informational sources or kinds of information, therefore, can undermine the effectiveness of individual decision-making. As Baekgaard and Serritzlew (2016) note, information loses its value if it is systematically misinterpreted. Furthermore, cynicism toward public sources of information may be linked to distrust and disengagement from important social institutions, undermining the health of our democracy.

This paper investigates whether individuals are biased against public organizations when faced with unambiguous performance information. We then explore if the credibility of the source of performance information varies across government regulators, nonprofit associations and private customer-ratings, and whether that credibility affects individual performance assessments. This study adds to the existing literature on anti-public sector bias in four ways. First, in addition to contributing to the evidence on bias in health care programs, this study

differentiates between for-profit and nonprofit organizations because many public services are delivered through both types of organization. Second, while studying respondents' perceptions of several traditional aspects of performance such as effectiveness, efficiency, equity, and red tape, we solicit respondents' views on whether they feel comfortable using organizational services. This approach offers a more holistic assessment signaling respondents' intentions with potential behavioral consequences. Third, while earlier studies have noted the critical role of goal ambiguity and its attendant informational asymmetries for individuals' capacity to perceive performance data, our study examines the effect of simple, commonly used performance measures – five-star ratings – to reflect good and bad performance. Fourth, while previous studies have not considered the credibility of information and its sources, our study examines its effect on individuals' performance assessments and thus contributes to the literature of credit attribution.

We perform our study using a randomized survey experiment with approximately 1,600 Americans using a vignette about a nursing home and its operations. Nursing homes are frequently examined in the public administration literature (Amrikhanyan et al. 2019; Jilke, Van Dooren, and Rys 2018), and the U.S. nursing home context is particularly interesting for assessing perceived performance. While historically, public nursing homes were linked to mid-century almshouses and infectious disease sanatoriums (Watson 2009), empirical research has affirmed the superiority of government and nonprofit homes compared to for-profit homes. Public and nonprofit homes have better equipment, more comfortable facilities, and resident control; they receive fewer complaints and have fewer regulatory deficiencies (Amirkhanyan, Kim, and Lambright 2008; Harrington et al. 2001; O'Neill et al. 2003; Riportella-Muller and Slesinger 1982; Santerre and Vernon 2005; Schlesinger and Gray 2006). While we do not expect

the general public to be fully abreast of these scientific findings, such differences in sector performance exist in many areas of public policy (Andrews et al. 2011; Hodge 2018).

## Public Perceptions of Government, Nonprofit, and For-profit Organizations

### Theories and Evidence on Anti-Government Bias

The contemporary views on government institutions in the U.S. are rooted in the ideas of independence, individualism, privacy, restricted government, and capitalism – the nation's core values explicitly documented in its historic documents. U.S. government spending as a portion of GDP is fairly limited compared to other Western democracies, and the overwhelming majority of Americans are employed in the private sector (Persson and Tabellini 1999). The general dissatisfaction and the lack of faith and trust in both the elected officials and executive agencies has long been a part of the American culture (Bok 2001; Goodsell 2015; Marvel 2015; Rölle 2017).

Some of these views can come from prior experiences (Hvidman and Andersen, 2016). These may involve personal negative experiences with public programs or bureaucrats, as well as direct observations of persistent unresolved social issues such as homelessness, poverty, crime, or food insecurity. Furthermore, perceptions of government as incompetent and wasteful can be rooted in the media's negative reporting on political scandals and incapable bureaucrats; they may also be linked to the key tenets of the New Public Management doctrine that stressed reduced scope of public programs and reliance on efficient, innovative and customer-oriented private entities; and, these perceptions may simply be related to a lack of knowledge about government and its work (Hvidman and Andersen 2016; Marvel 2016). The literature on individuals' attributions of blame suggests that public agencies are an easy target to blame when things go wrong (James et al. 2016; Marvel and Girth 2016) and that individuals often

misidentify private organizations as public when they are dissatisfied with services (Van Slyke and Roch 2004).

Such persistent negative views of government organizations are commonly referred to as "anti-public sector bias," defined by Marvel (2015, 210) as "(1) the expression of negative attitudes in the absence of supporting evidence—that is, evidence that public sector organizations or employees perform poorly; or (2) the expression of negative attitudes in the presence of countervailing evidence—that is, evidence that public sector organizations or employees perform well." The word "expression" repeatedly used in Marvel's definition underscores that the bias actively manifests itself through action: in a person's current decisions, opinions, and choices. What lies beneath the bias is a set of pre-existing long-term beliefs, assumptions and associations related to government institutions and government employees that may be unconscious, automatic, deeply engrained, and spontaneously activated (Hvidman 2018; Marvel 2016). The need to understand these perceptions is motivated by the fact that they can be harmful for the government, the public, and the society as a whole. These perceptions can discourage individuals from paying taxes, supporting public policies and programs, and complying with government regulations (Christensen and Lægreid 2005; Marvel 2015). They can also create a hostile environment for public employees, by lowering morale and hurting recruitment and retention, and reducing contributions from public co-production (Marvel 2015).

While many studies have explored negative perceptions of government, applying these negative perceptions to individuals' evaluations of performance is relatively recent. Several experimental studies randomized the stimuli, such as organizational ownership, and isolated its effect on respondents' perceptions of performance. Marvel (2015), for example, investigates the perceptions of a government agency (the U.S. Postal Service) as opposed to its private-sector

counterparts. He finds that respondents expected lower performance from the public organization, and this bias is not be overridden with positive information. Arguing that individuals' biases against the public sector are unconscious and automatic or implicit, Marvel (2016) further investigates anti-public sector bias using Implicit Association Tests to measure attitudes towards the public sector. This research on the postal service confirms that individual implicit attitudes weigh down evaluations of performance; and while countervailing performance information attenuates the bias, this informational effect is short lived.

An experiment by Hvidman and Andersen (2016) explores students' perceptions of public and private hospitals in Denmark in order to investigate the bias against public organizations. While all organizations were described to respondents as using modern technologies and having high levels of professionalism, public hospitals were perceived more negatively than private hospitals in terms of efficiency and red tape. Public hospitals received higher assessments on benevolence and lower ones on effectiveness albeit not significantly so. In another study, Hvidman (2019) investigates individuals' preexisting beliefs and how they influenced the perception of information about in-home elderly care. Despite identical performance information, public organizations were perceived as less effective. These assessments were conditional on respondents' predisposed beliefs about the public sector. Thus, the negative performance evaluations were only present among respondents who self-reported negative general attitudes towards the public sector. Meier, Johnson, and An (2019), however, replicate Hvidman and Andersen's (2016) hospital study in the U.S. twice (using both student and adult participants) and find no evidence of a bias against the public sector for any measure of performance.

Additionally, a randomized survey experiment by Baekgaard and Serritzlew (2016) investigates how Danish citizens interpret performance information about hospitals and schools based on their prior beliefs. Performance information involved the number of surgeries with complications and number of students who passed/failed a standardized test. They find that people are better able to interpret performance information correctly if it is consistent with their prior beliefs: those with more positive beliefs about the public sector were more likely to perceive positive information about it. Taken together, these findings suggest that additional research is needed to explore whether the bias exists across countries and policy fields.

The first objective of this study is to explore if the anti-public sector bias is present when assessing public and private organizations on separate dimensions of performance. Recognizing the unique role of the nonprofit sector in providing public services, we distinguish between private nonprofit and private for-profit organizations. Private nonprofit organizations are different from their for-profit counterparts even in health care, where all entities charge fees for services and face competitive pressures to attract clients, comply with government regulations, and generate revenues irrespective of their legal ownership. Prohibited from distributing their profits to owners or stakeholders, nonprofits tend to be more quality-oriented and trustworthy than their for-profit counterparts; they may be more flexible and prompt than governments in responding to urgent social issues, but may also suffer from unique forms of voluntary failure, such as particularism or amateurism (Amirkhanyan et al. 2018; Amirkhanyan, Kim and Lambright 2008; Eggleston and Zeckhauser 2002; Hansman 1996). Lumping private-nonprofit and private-for-profit organizations together, therefore, can result in missing important distinctions between these organizations.

**Hypothesis on Individual Perceptions of Outcomes across Sectors**

While comparing people's bias towards organizational ownership, this study focuses on several key dimensions of organizational performance: effectiveness, efficiency, red tape, equity, as well as respondents' comfort in using organizational services. Effectiveness generally refers to the "achievement of formal objectives" while efficiency is defined as "the ratio of outputs to inputs" (Andersen, Boesen, and Pedersen 2016, 853). Put differently, effectiveness is about producing intended results, while efficiency concerns optimally allocating resources to minimize waste. We expect respondents to attribute greater effectiveness and efficiency to private organizations – both nonprofit and for-profit – than governments for a number of reasons.

First, governments often are often portrayed as large and excessively formalized monopolies, ruled by "budget-maximizing bureaucrats" (Downs 1957; Downs 1967; Niskanen 1971). On the other hand, by virtue of operating in competitive markets with numerous other providers competing for clients, government contracts, donations and foundation grants, for-profit and nonprofit organizations can be view as striving to deliver the best services they can, as efficiently as possible, to satisfy their consumers and other stakeholders. Nonprofit organizations in particular, are pressured to perform well to retain their government contracts and to keep their administrative costs down for effective fundraising.

Second, by virtue of being subject to external and internal mandates, rules and procedures, government agencies are perceived as less flexible in hiring, firing, agency restructuring, long term-planning, and goal setting that could swiftly address emerging performance problems and stimulate innovation. Both nonprofit and private organizations have more flexibility in that regard and hence may be perceived as more flexible in correcting performance problems or innovating.

Third, the performance of public organizations is framed by the media and in the broader public and academic discourse as chronically poor: in the privatization literature, governments are characterized as generally bad at "rowing" and therefore are presenting a problem, that markets are able to solve (Savas 2000, 7). Nonprofit and for-profit sectors, despite the occasional performance scandals, overall do not share this negative reputation.

Fourth, while government agencies rely on public tax dollars to implement government programs, nonprofit service providers, in particular, can rely on private donations, service fees, foundation grants and volunteers and can therefore be perceived as more efficient and effective by better tailoring and connecting their deliverables to community resources, needs, and context. Empirical literature provides support that individuals tend to perceive public organizations as less effective than private organizations (Hvidman 2019). They are specifically more likely to pick for-profits when choosing a nursing home using ownership status in their search, unless they were better educated (Ben-Ner 2018). These studies and the arguments above suggest that private organizations – both nonprofit and for-profit – are more likely to be evaluated favorably in terms of effectiveness and efficiency.

Red tape indicates excessive unnecessary rules and formalities and the existence of administrative burdens (Keiser and Miller 2020). With the stricter personnel rules in governments, both academic research and media discourse suggest higher perceived levels of burdensome constraints in personnel and performance management in public organizations, compared to their private and nonprofit counterparts (Feeney and Rainey 2010). Consistent with this, we expect individuals to attribute higher levels of red tape to government agencies, as government actions are more likely to be bound by the many rules that help promote lawful and democratically accountable implementation of public policies. Nonprofit and for-profit

organizations, on the other hand, may be expected to have more flexible and less rigid internal environments. As a result, public organizations are likely to be rated higher in terms of red tape.

Social equity is another a critical value and goal in public administration. It concerns "fair and just treatment and the equal and equitable distribution of benefits to the society at large" (Riccucci 2009, 374; see also Frederickson 1971). Given the prominent role of government agencies and nonprofit organizations in serving lower socio-economic status individuals – serving as a safety net and providing food, shelter, and medical care to the most vulnerable social groups – individuals are likely to trust public and nonprofit organizations to emphasize fairness and social equity in service delivery. Existing research also shows that governmental agencies are perceived to be more equitable than private organizations (conceptualized broadly as a group) (Hvidman 2019).

Furthermore, both public agencies and nonprofit organizations do not distribute their profits to owners or stakeholders and are prohibited from paying unreasonably high compensation to its executives, unlike for-profit firms. In the health care setting, the non-distribution of profits reduces public and nonprofit organizations' incentives to pursue cost-cutting strategies and efficiency at the expense of serving marginalized clients; it also encourages them to perform better in terms of customer complaints and regulatory compliance (Amirkhanyan, Kim and Lambright, 2008; Eggleston and Zeckhauser, 2002).

The unique characteristics of public and nonprofit organizations can also shape citizens' willingness to personally rely on these organizations and receive services or benefits. Empirical evidence suggests that citizens tend to perceive nonprofits as warm, trustworthy, fair, and humane (Aaker et al. 2010; Drevs et al. 2014; Handy et al. 2010; Schlesinger et al. 2004). Meanwhile, they often associate for-profit forms and their workers with money or greed

(Willems 2020). In the context of hospitals, Drevs et al. (2014) find that trustworthiness is seen as higher in nonprofits and is an important dimension of patient evaluations.[2] We expect that these positive perceptions can lead to higher ratings in terms of comfort. In sum, public and nonprofit organizations would be rated higher in terms of equity and comfort in using services. Hence, we propose two hypotheses informed by the findings of prior research:

Hypothesis 1. Compared to for-profit and nonprofit nursing homes, public nursing homes will rate lower on effectiveness and efficiency and higher on red tape.

Hypothesis 2. Compared to for-profit nursing homes, public and nonprofit nursing homes will rate higher on equity and respondents' comfort with using their services.

## The Clarity and Credibility of Performance Information

### Source/Information Credibility

The principle of transparency – government's openness to public scrutiny – has been a key element of numerous government reforms (Hood and Heald 2006; Yang and Holzer 2006). The movement to collect, share, and make use of performance information, however, has been plagued with many problems: spamming individuals and bureaucrats with unsorted, complex or ambiguous information, encouraging administrative resistance, creating perverse incentives involving record manipulation, and other problems (Hood and Heald 2006; Moynihan 2006; Yang and Holzer 2006). Nonetheless, transparency of government actions and performance is an important accountability tool, a remedy for government deficiencies, and a prerequisite for better decisions (Hood and Heald, 2006; Moynihan, 2006). Americans have access to myriad informational sources about a given organization or a service, and whether the source of information matters in forming individuals' opinions, attitudes and behaviors has been raised (Pornpitakpan 2004). In particular, the credibility of the source and the performance information

was proposed as an important factor influencing how people view and use information. James and Van Ryzin (2017: 25) define credibility as "a combination of the perceived truthfulness of the level of performance reported on a measure and the perceived believability of the measure itself as a way of gauging performance." Likeability of the source, similarity to the audience, expertise and competence, and benevolent nature of the source have also been identified as important components of credibility for an informational source (Yoon, Kim, and Kim 2011).

To date, a few studies have addressed the credibility of performance information across sectors, and their findings are inconclusive. Van Ryzin and Lavena (2013) conduct a survey experiment about reporting street cleanliness information provided by the government, compared to a non-governmental or an unnamed source and found no significant difference in credibility by sector. Because the experiment involved an observable and easily measurable service, Van Ryzin and Lavena (2013) suggest that credibility effects might be more likely in a policy area with more information asymmetry and less tangible services so that the source may act as an actual cue. James and Van Ryzin (2017) focus on a less tangible service; they explore if people find customer satisfaction reported by government agencies to be more or less credible compared to an independent source. The findings show only a weak indication that the source is important with people viewing customer satisfaction as slightly less credible when the government agency reported about itself. The differences, however, are statistically insignificant. On the other hand, the political attractiveness of a government agency influences credibility, with agencies such as Citizenship and Immigration Services (CIS) having less credibility than agencies such as Veterans Affairs (VA). The difference in findings may also have resulted from the different levels of information asymmetry across policy areas.

**Hypothesis on Performance Information and Citizens' Perceptions**

To advance this line of inquiry, this article explores how clearly framed information combines with sector source and perceived credibility to influence assessments of organizational performance. Past studies that examined objective performance information and its effect on citizens' perceptions have generally provided context-specific data, such as student pass/fail rates, surgical complication rates, delivery rates and transit times (Baekgaard and Serritzlew 2016, Marvel 2015). Such information can be difficult to evaluate unless individuals are given data for comparative purposes (Olsen 2017b) since the different references levels of individuals can create ambiguity in terms of meaning (Holm 2017; Nagtegaal et al. 2020; Rothbart et al. 2019). This study seeks to provide less ambiguous data on performance: one that would preclude our respondents from wondering if the numeric indicators given should be viewed as acceptable, good or excellent.

This study focuses on the long-term care provided to senior or disabled clients who are unable to live in the community and perform activities of daily living. These services and their effects are not easily observable, and what constitutes "good" care may be subjective and conditional of the client's situation and other factors. Thus, rather than reporting the incidence of adverse outcomes, as was done in the past sector-bias studies, we will investigate the effect of information presented using a simple five-star quality rating system. As detailed in Methods, this system is used in nursing home care, presenting not only a simple tool for differentiating performance across organizations, but also a realistic approach currently used to inform customers about nursing home care quality.

Hypothesis 3. Positive performance information provided to respondents will lead to higher perceived effectiveness, efficiency, equity and comfort, and lower red tape.

**Hypotheses on Source Credibility and Perceptions of Performance across Sectors**

While the nature of information matters, its effect on perceptions will likely be moderated by how credible the source and its performance assessment practices are. More specifically, we focus on whether source and assessment credibility vary by public, nonprofit and for-profit sector. When it comes to the sector implementing these assessments, the absence of the profit motive, independence, and regulatory and professional expertise of government and nonprofit information sources should serve as signals of credibility, resulting in higher public ratings compared to the ratings of for-profit information sources. For-profit evaluators may not be fully independent of the organizations that are assessed, the performance metrics that are used, how the entities are rated, and who participates in the assessment. As an example, in 2015, a class action suit was filed in federal court alleging that a well-known business assessment website, Angie's List, manipulated companies' ratings to favor its advertising revenues (Fiorillo 2015). The absence of the profit motive, therefore, may result in higher levels of trust towards public and nonprofit sources.

Similarly, public and nonprofit organizations, as noted earlier, are extensively regulated and are subject to transparency requirements to justify their tax-exempt status and to promote their accountability to the taxpayers, legislature, interest groups, media, and other parties. Nonprofit organizations, in particular, are subject of oversight for numerous institutions. Besides government regulators and funders, these include professional associations that provide accreditation and certifications needed to compete and signal quality, as well as sector-specific national and local charity-rating organizations such as *Charitynavigator.org*, *Charitywatch.org*, *Candid.org* (a new organization that united former *Foundation Center* and *GuideStar.org*). The latter group provides detailed information about nonprofit organizations to potential donors,

researchers, partners and other constituents. These pressures are likely to lower the moral hazard among public and nonprofit evaluators.

While it is relatively easy to experimentally manipulate organizational ownership in a multi-sector context by assigning public, nonprofit or for-profit status to identical organizations, the task of manipulating the ownership of informational source while maintaining a realistic assessment scenario is more challenging. In this study, we seek to explore the effect of sector-specific assessments that more closely reflect real-world practices used by public, nonprofit and for-profit entities conducting the assessment. Government assessments and ratings aiming at Medicare or Medicaid funded hospitals, nursing homes, or Head Start childcare providers, are commonly implemented by regulatory bodies (teams of experts) that follow a standard formal protocol involving a process to identify violations of a predetermined set of regulations (Health Deficiencies n.d.; Amirkhanyan, Kim and Lambright 2010). Thus, the process for public-sector assessments more commonly involves expert assessment of compliance with a set of objectively and formally stated criteria.

Private sector assessments, on the other hand, rely more heavily on customer-reviews or assessments by professional groups. Thus, nonprofit accrediting bodies, often tie the entities' accredited status to these entities' formal membership which may also entail annual fees. For-profit sources – such as Amazon or Google – on the other hand, often involve perceptual customer-ratings. Thus, in this current study, we explore the credibility for a *combination of both the sector of the party conducting the assessment and the process for this assessment*. As detailed in the methods section, we differentiate between a government agency that inspects and assesses nursing homes in accordance with a federal mandate; a for-profit web platform that aggregates customer ratings of nursing homes, and a nonprofit association that conducts accreditation

reviews and evaluates nursing facilities. Accordingly, this article proposes two additional hypotheses:

Hypothesis 4. Government inspectors and nonprofit accrediting agencies, as sources of information, will rate higher on credibility, compared to for-profit customer-driven sources of information.

Hypothesis 5. High credibility will enhance the effect of objective performance information on all perceived measures of organizational performance.

## Methodology

### The Research Design

This study tests the hypotheses using an on-line survey experiment of 1,600 US respondents through Amazon Mechanical Turk (MTurk). The number of respondents was determined by a power analysis using the conventional threshold, statistical power of .80 at a significance level of .05 (Walker et al. 2019). Prior investigations have shown that the average MTurk user is more likely than the population to be female, more educated, and lower income (Paolacci and Chandler 2014; but on gender see Jilke et al. 2016; Marvel 2015; 2016). Less than half of MTurk participants work full-time for an organization (Stritch, Pedersen, and Taggart 2017). Hispanic and Black respondents are underrepresented while Asian respondents are overrepresented (Berinsky et al. 2012). Even if MTurk generates samples that are less representative of the US population, the randomized nature of the experiment and existence of possible controls should be able to adjust for any bias. Recent research demonstrates that MTurk subjects are more representative of the general population than other convenience samples used by researchers, therefore, they can produce high-quality data (Berinsky, Huber, and Lenz 2012; Buhrmester, Kwang, and Gosling 2011).

While MTurk provides access to a more representative sample, respondents on this, and similar, platforms have a financial incentive to complete surveys quickly and multiple times (Stritch, Pedersen, and Taggart 2017). To address these limitations and ensure the quality of our data, we incorporate several features into our design. First, on MTurk, we limited the participation of the survey to US residents. Second, we eliminated respondents who spent less than 60 seconds on the survey (16 total respondents). Rushing through the survey can negatively impact reliability of the data and increase the likelihood of errors (Button et al. 2013). Creating a threshold for time to complete the survey and rejecting those who do not meet the threshold is one of many ways to address this concern (Stritch, Pedersen, and Taggart 2017). Third, we eliminated respondents who answered the survey more than once. Additionally, we conducted series of robustness checks detailed below, including those involving various time frames for survey completion, as well as those determining if people recognized the sector, the rater and the star-rating they were given.

Our survey experiment uses a between subjects $3 \times 3 \times 3$ factor design. At stage 1, respondents are randomly assigned to evaluate a fictitious government, private nonprofit, or private for-profit nursing home.

> **Group 1 read:** "Greenfield Meadows is a government owned nursing home. It has been owned and operated by the county government since 1999. The Nursing Home Administrator of Greenfield Meadows reports directly to the County Executive."
>
> **Group 2 read:** "Greenfield Meadows is a private nonprofit nursing home. It has been owned and operated by the Greenfield Coalition for Senior Care, a state-wide nonprofit organization since 1999. The Nursing Home Administrator of Greenfield Meadows

reports directly to the nonprofit Board of Trustees of the Greenfield Coalition for Senior Care."

**Group 3 read:** "Greenfield Meadows is a private for-profit nursing home. It has been owned and operated by the Greenfield Senior Care Corporation, a for-profit long-term care company since 1999. The Nursing Home Administrator of Greenfield Meadows reports directly to the Chief Executive Officer of the Greenfield Senior Care Corporation."

The description is accompanied by a drawing of a two story-building with a "Greenfield Meadows" sign on its front lawn. Then, *all* respondents read the following vignette:

"The nursing home has 85 full-time employees and an average of 94 residents on a given day. The nursing team consists of state licensed registered nurses, practical nurses, and geriatric nursing assistants supervised by a Director of Nursing. Every nursing home resident gets, on average, about 4 hours of direct nursing care per day. Every resident is under the care of either the Medical Director or their own physician. Greenfield Meadows' goal is to provide long-term care of the highest quality, taking into account the special needs of the individual residents. In addition to long-term care, Greenfield Meadows offers skilled nursing care, memory care, rehabilitation, respite care, hospice and palliative care, social services, wellness programs, and diverse social, educational, and recreational activities."

At stage 2, the respondents are randomly assigned to an outside evaluator who rates the nursing home. CMS is the official federal government unit that coordinates the evaluation of nursing homes. NHAA is a fictitious organization. Google Health does not exist, but one can find such ratings of nursing homes on-line sponsored by Google.

> "Centers for Medicare and Medicaid Services is a federal government agency that administers the Medicare and Medicaid programs and works in partnership with state governments to regularly inspect and evaluate nursing homes on a scale from 1 to 5 stars."
>
> "Nursing Home Association of America is a private nonprofit association of nursing homes working on advancing nursing home care through education, research, leadership, and practice. Based on annual accreditation reviews, Nursing Home Association of America evaluates nursing homes and other long-term care organizations on a scale from 1 to 5 stars."
>
> "Google Health Care is a private for-profit organization that uses an online platform to aggregate customer reviews about health and nursing care providers on a scale from 1 to 5 stars."

At stage 3, respondents are told that the evaluator has rated the nursing home as two stars, three stars or four stars on a five-star rating scale. The star rating should produce a relatively clear signal and has mundane realism. The official performance ratings for both nursing homes and hospitals in the US use a five-star system as do many commercial rating systems such as Yelp or TripAdvisor. Other systems that do not use "stars" per se, such as those in K12 education, use similar ordinal scales for ranking.

Following exposure to the experimental conditions, respondents were asked to evaluate the nursing home based on several performance dimensions. Respondents were then asked a series of demographic questions, followed by items to measure pre-existing sector bias used by Hvidman and Andersen (2016), whether they or any of their family members have experience with nursing homes, and an assessment how credible the performance evaluations were

perceived. We then asked manipulation checks for nursing home sector, evaluator organizations, and the actual rating.

Because this study involved one or more fictitious organizations (depending on what scenario a respondent was randomly assigned to), at the end of the survey, all respondents were informed of the study objectives and whether the organizations they read about were real (e.g., CMS) or fictitious. The design and the instruments of this study were reviewed and approved by the American University Institutional Review Board.

**Dependent Variables**

Public programs can be evaluated on a wide variety of dimensions including effectiveness, efficiency, equity, and red tape (Boyne 2002; Hvidman and Andersen 2016; Meier et al. 2019). For each of the concepts, individuals were asked to rate the nursing home on a set of indicators with ratings on a seven-point scale with "fits very well" at one end of the scale and "does not fit at all" at the other end of the scale (see Table 1). All scales were constructed via principal components using the single significant first factor (Hall and Van Ryzin 2019). The efficiency scale uses four indicators that all load at 0.81 or better with a Cronbach's alpha of 0.90; the current scale is significantly more reliable that existing scales used in the literature. Effectiveness is measured with eight indicators all loading at 0.89 or better with a Cronbach's alpha of 0.97. The red tape scale is based on two items used in the literature (see Hvidman and Andersen, 2016) with loadings at 0.89 and a Cronbach's alpha of 0.73. Equity is based on three indicators each loading at 0.88 or better with a Cronbach's alpha of 0.88.

Finally, we report the mean and the standard deviation for our measure of "comfort" ("If a member of your family was in need of nursing home care, how comfortable would you be placing your relative in Greenfield Meadows?"). While other dependent variables explicitly

focus on specific aspects of organizational performance, this measure reflects a broader, more holistic attitudes reflecting one's propensity to engage with an entity. One could feasibly perceive an organization as inefficient, have valid concerns about the inequity resulting from its work, or have concerns about the overall ineffectiveness of its actions in terms of mission or goal attainment. Despite this, someone can still be fully comfortable receiving benefits and using services delivered by an agency. The R-Square for "comfort" explained by other dependent variables ranges from 0.352 to 0.566 (for efficiency, effectiveness, and equity). Thus, individual comfort in using services may be determined by a range of other factors: personal need, ideology, value given to equity or efficiency, and others. Empirically, all the experimental findings in terms of "comfort" remain even when the four other performance indicators are included in the equations indicating the comfort picks up performance dimensions separate from effectiveness, efficiency, equity, and red tape (results available from the authors). Although our survey design does now allow us to observe respondents' actions, this measure can effectively signal their intentions more than other dependent variables used in the current study or past research on sector bias.

[Table 1 here]

**Balance Tests**

To check on the randomization of the $3 \times 3 \times 3$ experiment, balance tests were run based on the respondents' age, education, income, and partisanship relative to the experimental conditions of sector, evaluator, and star rating. In none of the twelve cases do the distributions deviate statistically from random based on the f-tests tests (see Table 2); the randomization resulted in balanced groups.

[Table 2 here]

## Findings

In Table 3, our OLS regression results present the relative evaluations of public, nonprofit, and for-profit nursing homes with each of the five dependent variables – efficiency, effectiveness, red tape, equity, and comfort.[3] The reference category is private for-profit nursing homes. The constant is the mean value on each of these variables for for-profit nursing homes; and the values for government and nonprofit reflect how different those homes are rated relative to for-profit homes. As an example, in terms of efficiency, for-profit homes are rated at -0.060 or effectively neutral (the dependent variables other than comfort are factor scores with a mean of 0 and a standard deviation of 1), given that the constant is not statistically significant. The government coefficient indicates that on average government operated nursing homes are rated 0.056 points higher on this scale or at -0.004. The differences are not statistically significant, indicating that the public does not perceive government nursing homes any different from for-profit nursing homes. In short, this indicates a lack of public sector bias. The nonprofit coefficient, however, is statistically significant indicating that nonprofit nursing homes are perceived as *more efficient* than for-profit homes. At the same time, this difference is small (0.125), a fraction more than one-tenth of a standard deviation. A comparison of the coefficients and their standard deviations also indicates that the respondents do not perceive any difference between nonprofit and government nursing homes on this dimension.

[Table 3 here]

The findings for efficiency are essentially replicated for effectiveness and equity. In both cases individuals also see nonprofit nursing homes as performing slightly better (more effective, more equitable) than for-profit homes, and the respondents see no difference between government nursing homes and for-profit homes. The magnitude of the differences is modest. On

the red-tape dimension, all coefficients are statistically insignificant, and there are no perceived differences across any of sectors. Finally, on the comfort in placing a relative in the nursing home, the coefficient for nonprofit facilities is 0.230 and statistically significant, indicating that respondents were approximately a fifth of a standard deviation more comfortable placing a relative into a nonprofit nursing home than a for-profit home.

Although the difference is not as great, respondents also feel more comfortable with nonprofit nursing homes than public ones. When we set public homes as the reference group (to test whether there is a significant difference between public and nonprofit facilities), the nonprofit coefficient is statistically significant in the comfort model at a modest level ($b = 0.145$, $p < 0.065$). Public nursing homes, on the other hand, did not differ from the for-profit homes in all aspects of performance (see Table A1 in the appendix) suggesting that for-profit performance overall is not rated any better than government performance. In summary, we find no support for hypothesis 1 and only partial support for hypothesis 2, as originally stated. There is no evidence of anti-government bias, and respondents rate for-profit nursing homes lower on efficiency, effectiveness, equity and comfort of use compared to nonprofit homes.

The sector bias literature also proposes that public organizations do not get credit for their performance relative to private organizations (Marvel 2015), although in some cases no organizations get credit for performance regardless of sector (Hvidman and Andersen, 2016; Meier et al. 2019). This study hypothesizes that the credibility of the source of the performance evaluation might be a factor as well as the ambiguity of the performance evaluation. To deal with these concerns, the performance evaluation was changed to a highly common government performance system (a five-star rating system), that is actually an official performance rating system for nursing homes in the U.S. The experiment randomly assigned two, three, or four stars

to the nursing home. The credibility of the rating source was varied between the Centers for Medicare and Medicaid Services (highly credible), the Nursing Home Association of America (credible) and Google Health (less credible).

Table 4 presents one look at the impact of the performance evaluations and the credibility of the source. The star rating is the measure of performance presented to the subjects. The table also interacts the star rating with the source of the evaluation either CMS or NHAA; Google Health is the reference category. Our findings support Hypothesis 3: for four of the dependent variables – efficiency, effectiveness, equity, and comfort – the star rating is positively and strongly related to the performance of nursing homes; the influence is at least one-half standard deviation for an additional star. For red tape, the star performance generates a small negative significant relationship. Although nothing in the experimental vignette discussed red tape, this suggests a halo effect on different performance dimensions in that well performing nursing homes are also perceived as having fewer burdensome regulations.

[Table 4 here]

A walk-through of the ratings will make them a bit more interpretable. Nursing homes could be rated as 2, 3 or 4 stars. These ratings need to be calculated with respect to the constant reported in the tables. The average efficiency rating of a nursing home getting four stars (ignoring the insignificant coefficients) is $-1.678 + (4 \times 0.568)$ or 0.594, about a half of standard deviation above the mean. The mean for three stars is 0.026 or essentially 0, the overall midpoint of the evaluation, and -0.542 for a home getting only two stars. Similar calculations can be made for the other dimensions.

Table 5 interacts the star rating with the sector of the nursing home either government or nonprofit; for-profit is the reference group. Again, the star rating is positively and significantly

associated with the performance of nursing homes. None of the interaction terms between sector and star ratings indicating whether or not respondents discount performance information in either public or nonprofit sectors are statistically significant.

[Table 5 here]

The results in Tables 4 and 5 treat the star rating system as an interval variable even though much of the literature suggests that negative information is more important than positive information (e.g., Charbonneau and Ryzin 2015; Hong and Kim 2019; James et al. 2016; Nielsen and Moynihan 2017). To determine if these results are misleading due to negativity bias, Table 6 estimates the impact of the star rating system by using a dummy variable for those rated 4 stars and a second dummy variable for those rated 2 stars; the reference category is the homes rated 3 stars. If there is a negativity bias in the response to the rating system, the absolute value of the coefficients for two stars would be much larger (although negative) than the absolute value of the coefficients for four stars. While each of the two-star coefficients is slightly larger in magnitude than the four-star coefficients, the differences are trivial. The clear conclusion is that in this case, respondents do not appear to incorporate negativity bias into the five-star rating system.

[Table 6 here]

Returning to Tables 4 and 5, the clear positive influence of the rating system stands in contrast to the results in regard to sector and the rating source. None of the ten interaction terms linking who did the rating with the actual rating are statistically significant (Table 4). This result suggests that the evaluator of nursing homes was irrelevant, that the rating does not matter either more or less if given by a different source. In addition, none of the sector variables are statistically significant (Table 5). *When individuals are given performance information on nursing homes, they rate nursing homes solely on the star rating given and do not rate*

*government nursing homes any different from either for-profit or nonprofit nursing homes*. For government nursing homes the coefficients are all in the direction of government homes getting more credit (or having less red tape), but those results are not statistically different from zero.

Before concluding that the credibility of the evaluator plays no role in the assessment of public programs, some additional analysis is merited. The experiment also asked respondents questions about how credible and trustworthy they thought the rating organization was. Specifically, the questions included (1) whether they thought the star rating assigned to the nursing home was (definitely) true or not and (2) how believable the star rating as a way of evaluating nursing home care was (see Table 1). Using principal component analysis involving these two survey items, a scale was constructed and used in regression analysis. Table 7 shows how the assessment of credibility is in fact related to who the evaluator was. Supporting Hypothesis 4, our findings indicate that respondents whose ratings came from the CMS or the NHAA perceived their information source as significantly more credible than those whose ratings came from Google Health. The CMS and NHAA credibility ratings, on the other hand, were not statistically different from each other.

[Table 7 here]

In Table 8, we interact how credible the individual found the evaluator with the star rating and find that if the individual perceived that the evaluator was more credible, then this positively contributed to the overall evaluation. A one standard deviation increase in perceived credibility added between 27% and 88% to the impact of the star rating (efficiency + 46%, effectiveness + 46, red tape + 88%, equity + 68% and comfort + 27%). Supporting Hypothesis 5, perceived credibility clearly matters in how performance information is interpreted. Including the credibility interaction also brings back the bias in favor of nonprofits (relative to for-profit

nursing homes) in terms of comfort. In addition, there is a slight bias in favor of government nursing homes relative to for-profit homes in terms of equity that does not reach traditional significance levels. These differences are small and unlikely to be of any substantive importance in the evaluation of nursing homes. In addition, there are no statistically significant differences between government and nonprofit nursing homes. It would make sense to characterize the modest biases that were found as anti-for-profit biases in this industry.

[Table 8 here]

**Robustness checks**

Additional manipulation and robustness checks were performed and are available from the authors. Specifically, we obtained tables 3 through 7 for a variety of sub-groups: (1) the sector match sample (those who correctly reported the nursing home's sector information at the end of the survey), (2) the evaluator match sample (those who correctly identified the evaluator at the end of the survey), (3) the star-rating match sample (those who correctly identified the star ratings at the end of the survey), (4) the sector and star-rating match sample, (5) the sector, star-rating, and evaluator match sample,[4] (6) the public-sector preference subsample, which includes those who showed preference for public organizations, (7) the private-sector preference subsample, which included those who showed preference for private organizations, (8) the Democrat subsample, (9) the Republican subsample, (10) the "independent/no party affiliation/other" subsample, (11) the 5+ minutes subsample (those who took over 5 minutes to complete the survey), and (12) the <5 minutes subsample. None of these additional analyses generated results that were substantively different from those presented here.

Two of these robustness checks merit comment. By examining the results both by those who favored public sector (versus private sector) and those who identified as Democrats versus

Republicans, we can reject the notion of motivated reasoning in this particular case. Pre-existing biases did not result in individuals' discounting performance information or in rating the nursing homes differently. A final robustness check involved examining those who stated that the respondent or their family member had experience with using a nursing home, and in this case, there was a slightly stronger bias in favor of nonprofit nursing homes (see supplemental appendix).

## Discussion

Within the broader category of health and human services, nursing home care is a vital service for persons with disabilities, regardless of race, income, and social status. In the context of this heavily regulated and, largely, publicly funded industry, nursing home performance information has been available to citizens for over three decades: first, using staffing ratios and counts of regulatory violations and, since 2008, using facility five-star ratings available for all public, nonprofit, and for-profit Medicare or Medicaid certified nursing homes (Abt Associates Inc., 2014). Understanding whether individuals are biased against one sector as a provider of care or a source of performance information and method of assessment has implications for the use of these facilities and government's efforts to be transparent on their performance.

We examine individuals' perceptions of separate performance dimensions for public, nonprofit, and for-profit nursing homes, using an online-survey experiments via MTurk. While there are concerns regarding data quality collected via Mturk, such as sample representativeness and the small population of MTurk workers (Paolacci and Chandler 2014), this engine has provided reliable and useful data for researchers (Berinsky, Huber, and Lenz 2012; Buhrmester, Kwang, and Gosling 2011). *Our analysis finds little evidence of anti-public sector bias. In no case were private for-profit nursing homes rated significantly better than government nursing*

*homes*. However, on several dimensions – efficiency, effectiveness, equitability, and comfort – nonprofit nursing homes are perceived more favorably than for-profit nursing homes. Thus, there is some evidence of anti-for-profit sector bias in the field of nursing home care, providing a caution against grouping for-profit and nonprofit providers together in assessments of public service delivery.

The lack of anti-government bias we find is consistent with Meier et al. (2019) findings on U.S. hospitals. The latter study began with a theoretical expectation that the biases would be more pronounced in the U.S. than in Denmark (Hvidman and Andersen, 2016) since anti-government sentiments and bureaucrat bashing are more prevalent in the U.S. Given their null findings, Meier et al. (2019) suggest that in a mixed delivery system, stereotypes may in fact be weaker and less relevant because the sector is not always evident to the client. Similarly, in nursing home care, all facilities irrespective of ownership are subject to licensure and certification requirements; they also admit publicly funded residents and get reimbursed by government funds. They all strive to remain competitive and invest in their appearance. Similar to the hospitals investigated by Meier and colleagues, nursing home care as an industry may be more homogeneous and more "public" as a field than other policy areas where public-private distinctions may be more pronounced and therefore associated with a bias.

A possible explanation for our findings may be related to the fact that performance data on nursing homes (via Nursing Home Compare) has been publicly available since 1990s and over a million Americans rely on nursing home care at any given time (National Center for Health Statistics 2019). While early online versions of Nursing Home Compare did not allow clients to compare groups of nursing homes by sector, they clearly presented deficiencies for all nursing facilities in a given geographic area (e.g., a county). With consistent scientific evidence

on the inferiority of for-profit nursing homes (Amirkhanyan, Kim, and Lambright 2008; Harrington et al. 2001; O'Neill et al. 2003; Riportella-Muller and Slesinger 1982; Santerre and Vernon 2005; Schlesinger and Gray 2006), it is possible that individuals' perceptions may be a function of their personal experiences and knowledge of how facilities compare across sectors. We cannot, therefore, rule out that the anti-for-profit bias in our study may reflect some level of consumer knowledge in this field. This is supported by one of our robustness checks: respondents with prior personal/family nursing home care exposure or experience had a slightly stronger bias in favor of nonprofit facilities.

One important implication of our research is the need to explicitly differentiate between private-for-profit and private-nonprofit organizations. Our study suggests these may be perceived differently in some policy fields. Particularly in the United States but also in a variety of other countries, public programs are implemented by nonprofit organizations in health care, social services and other policy areas. Explicitly separating these organizations into distinct categories will be informative. In addition to the cross-sector differences on the traditionally used measures of organizational performance, our findings with respect to the new measure of individual comfort ("how comfortable are you using this organization for a relative?") are also significant. Thus, while the science of organizations differentiates among numerous dimensions of performance, individuals may more easily share their preferences through more holistic measures reflecting their intentions.

A second key finding pertains to the substance and credibility of performance information presented to respondents. In the era of "fake news" and declining citizens' trust in government institutions, understanding what signals can be associated with greater citizens' approval and cooperation is critical for the effective implementation of public policies and

programs particularly in the democratic context. Our study finds that performance information plays a critical role in how people view every aspect of organizational performance. Respondents rate nursing homes solely on the ratings they received and do not assess organizations from one sector any differently than other organizations. Diverging from the past experiments in hospital care, our study finds no evidence of people discounting good news and focusing on bad news. This may be because our experiment uses simple and clear performance indicators (five-star ratings) that are in fact commonly used in rating public programs. Overall, this finding suggests that citizens are able to learn and change their opinions based on performance data they are exposed to, and the latter is true for successes and failures of not only private but also public organizations. Additionally, the fact that a third party (a named organization) was charged with designating these ratings may have suggested to our respondents some degree of objectivity.

While it is no surprise, perhaps, that survey participants took performance data into consideration, the data on the American citizens' low and declining trust in government institutions raises questions about the credibility of performance data originating specifically from and supplied by government institutions or government experts. The long-standing "bureaucrat bashing" culture may have extended to citizens views related to the program evaluation, monitoring and oversight functions of government. Notably, however, when assessing the credibility of the source, our survey respondents rated the government agency (CMS) and the nonprofit association (NHAA) significantly higher than its for-profit counterpart (Google Health). Thus, who delivers the news and who conducts assessments is important. Here again, our findings may suggest some level of anti-for-profit bias in people's assessments of an evaluator and the process of evaluation. These findings may reflect our respondents' general preference for nonprofit and public organizations compared to for-profit organizations.

Government and nonprofit watchdogs may be seen as more "mission-driven," "impartial" or "neutral," and less likely to financially benefit from their assessments of other entities. While realizing that regulation involves time and money, the public may be more likely to expect and demand more from government agencies in terms regulating and enforcing higher standards on the quality of goods and services they are consuming. Additionally, assessments conducted by experts – government inspectors or nonprofit accrediting agencies – are assigned higher levels of credibility than consumer-generated performance data. When such assessments are provided, our experiment suggests that they are indeed taken more seriously and seen as more credible than assessments conducted and shared through the for-profit channels.

It is worth highlighting that Google Health was the only online and customer-driven evaluator we presented to our respondents, while the CMS and the NHAA quality assessments were done by these organizations and their staff. Lower credibility might have been linked specifically to consumer-driven assessments (Google Health Care), compared to inspection or accreditation-based assessments (CMS and NHAA), which may have been presumed by our respondents to be more systematic, following a standard protocol, focusing on key aspects of care, and conducted by experts. Follow-up research may be needed to explore if sector matters when all three sectors follow identical oversight and evaluation procedures, or when all three sectors conduct consumer-driven assessments. In the case of our paper, we opted for slightly different descriptions in order to better tie the evaluation to the real-life options currently available in nursing home care.

## Conclusion

Recent research in public administration has applied psychological theories to better understand citizens' perceptions of government performance (e.g., Hvidman 2019; Olsen,

2017a). These studies examine how individuals interpret performance information and whether there is bias against public organizations in the process (e.g., Hvidman and Andersen 2016; Marvel 2015; 2016). While these studies shed light on the psychological process of individuals' evaluation of government, less emphasis has been placed on perceptions of performance information credibility and the appropriate comparisons (nonprofits as well as for-profit comparisons). To advance this line of research and extend it to issues of public management, this article investigates whether individuals are biased negatively in their evaluation of public organizations in the context of U.S. nursing homes, and whether performance information from government agencies is perceived as less credible compared to private sources.

The results of our experiment have positive implications for issues of democratic accountability in public programs that are implemented via public, for-profit or nonprofit organizations. Members of the public are able to use the type of performance information that is frequently provided by government to assess the quality of nursing home services. They also incorporate information in more nuanced ways than previously found in the literature; they exhibit no negativity bias, and they emphasized information from sources they thought were more credible. The anti-private biases that did influence perceptions were consistent with existing objective assessments of performance and, thus, provided a rationale for discounting reported performance. Elder care provision is not a policy area/market that makes it easy for consumers to exercise informed judgement. It is characterized by extensive information asymmetry, the existence of third-party payers, and high transactions costs associated with finding or changing a nursing home. Even with these structural difficulties, however, respondents were able to make informed assessments of service quality. These findings suggest

that the public may be able to make similar quality judgements in other policy areas if they are provided with clear, unambiguous indicators of performance from credible sources.

## Notes

[1]Here and elsewhere in the paper we use the term "citizen" broadly, to denote all the people residing in a country and/or state/local jurisdiction.

[2]While Handy (2010) also demonstrates that nonprofits are seen as more trustworthy, people are also less able to identify nonprofits, illustrating that people are less familiar with this sector. Similarly, Schlesinger et al. (2004), similarly finds that, generally, the public does not understand organizational ownership well and tends to hold more negative perceptions of nonprofit health care performance than for-profits. These perceptions remain until ownership is explained, in which case perceived nonprofit legitimacy increases (Schlesinger et al. 2004).

[3]With five dependent variables and multiple tests, the power of statistical tests weakens and can generate false positives. The Holm-Bonferroni technique (full results available from the authors) indicates only a small number of possible false positives (two of the five anti-for-profit bias relationships; one of the 15 star relationships; one of the ten two and four star relationships; and none of the ten credibility or credibility interaction relationships).

[4]Of the respondents 61% passed the manipulation check for sector, 79% for the evaluator, and 84% for the star rating. Although 61% might seem low for a manipulation check, a recent study of people who actually selected a nursing home for a family member found between 33% and 44% (depending on how those not responding are coded) could correctly identify the sector of the home (Ben-Ner, Hamman, and Ren, 2019). For the sector cue, individuals were told the sector of the nursing home twice and were also told who the nursing home director reported to as a third sector indicator. More importantly, the significance tests for the manipulation check indicate that the treatment was sufficiently strong to elicit a response.

# Tables

**Table 1. Factor-Analytical Results of Survey Items**

| Survey Item | Factor Loading |
|---|---|
| **Efficiency** | |
| The nursing home provides care efficiently. | 0.89 |
| The nursing home makes the most of its monetary and human resources. | 0.88 |
| The nursing home is not wasteful. | 0.81 |
| The nursing home resources are well spent. | 0.91 |
| Eigenvalue = 3.05 | |
| Cronbach's alpha = 0.90 | |
| **Effectiveness** | |
| The nursing home is effective. | 0.91 |
| The nursing home is effective in accomplishing its core mission. | 0.91 |
| The nursing home is effective in delivering a very good service. | 0.92 |
| The nursing home is genuinely interested in the well-being of its residents. | 0.89 |
| The nursing home acts in the interest of residents. | 0.90 |
| The nursing provides outstanding quality of care. | 0.89 |
| The nursing home ensures excellent quality of life for its residents. | 0.92 |
| The nursing home ensures a clean, safe, home-like and comfortable environment. | 0.89 |
| Eigenvalue = 6.53 | |
| Cronbach's alpha = 0.97 | |
| **Red Tape** | |
| The nursing home has a high level of burdensome administrative rules and procedures. | 0.89 |
| A high level of administrative procedures negatively affects the nursing home's effectiveness. | 0.89 |
| Eigenvalue = 1.58 | |
| Cronbach's alpha = 0.73 | |
| **Equity** | |
| The nursing home delivers care to residents in a fair and impartial way. | 0.89 |
| Every resident, regardless of race, religion or income, gets the same quality of care. | 0.93 |
| Persons of any race, religion or income have an equal chance of being admitted to this nursing home. | 0.88 |
| Eigenvalue = 2.43 | |
| Cronbach's alpha = 0.88 | |
| **Comfortable** | |
| If a member of your family was in need of nursing home care, how comfortable would you be placing you relative in Greenfield Meadows? | |
| (5-point scale from "very comfortable" = 5 to "very uncomfortable" = 1) | |
| Mean = 3.30, SD = 1.29 | |
| **Credibility** | |

| | |
|---|---|
| Do you think the star rating assigned to the nursing home by [evaluator cue] is..." (5-point scale from “Definitely true” = 5 to “Definitely not true” = 1) | 0.89 |
| In your view, how believable is the star rating as a way of evaluating nursing home care at Greenfield Meadows? (10-point scale from “Completely believable” = 10 to “Not believable at all” = 1) | 0.89 |
| Eigenvalue = 1.57 | |
| Cronbach’s alpha = 0.59 | |

**Table 2. Balance Test Results**

| | | Sector | Evaluator | Rating |
|---|---|---|---|---|
| Age | F | 0.42 | 1.98 | 0.64 |
| | Prob > F | 0.66 | 0.14 | 0.53 |
| | N | 1,599 | 1,599 | 1,599 |
| Education | F | 1.97 | 0.11 | 1.95 |
| | Prob > F | 0.14 | 0.89 | 0.14 |
| | N | 1,607 | 1,607 | 1,607 |
| Income | F | 0.21 | 1.31 | 0.47 |
| | Prob > F | 0.81 | 0.27 | 0.63 |
| | N | 1,606 | 1,606 | 1,606 |
| Partisanship | F | 0.69 | 0.36 | 0.99 |
| | Prob > F | 0.50 | 0.70 | 0.37 |
| | N | 1,599 | 1,599 | 1,599 |

**Table 3. Overall Sector Bias**

| | Efficiency | Effective | Red tape | Equity | Comfortable |
|---|---|---|---|---|---|
| Government | 0.056 | 0.048 | -0.020 | 0.100 | 0.085 |
| | (0.063) | (0.063) | (0.061) | (0.062) | (0.080) |
| Nonprofit | 0.125* | 0.149* | -0.050 | 0.129* | 0.230** |
| | (0.061) | (0.062) | (0.062) | (0.061) | (0.079) |
| Constant | -0.060 | -0.066 | 0.023 | -0.076 | 3.196*** |
| | (0.044) | (0.045) | (0.043) | (0.045) | (0.057) |
| R-squared | 0.003 | 0.004 | 0.000 | 0.003 | 0.005 |
| *N* | 1,578 | 1,558 | 1,601 | 1,600 | 1,607 |

Note: OLS regression. Standard errors are shown in parentheses. The reference group for public and nonprofit nursing homes is for-profit homes. Two-tailed tests, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

**Table 4. The Credibility of the Source Star Ratings**

| | Efficiency | Effective | Red tape | Equity | Comfortable |
|---|---|---|---|---|---|
| CMS | -0.159 | 0.031 | 0.179 | 0.091 | 0.272 |
| | (0.209) | (0.200) | (0.234) | (0.218) | (0.249) |
| NHAA | -0.062 | -0.026 | 0.332 | 0.109 | 0.346 |
| | (0.210) | (0.201) | (0.234) | (0.219) | (0.250) |
| Star-rating | 0.568*** | 0.678*** | -0.122* | 0.485*** | 0.982*** |
| | (0.050) | (0.047) | (0.055) | (0.052) | (0.059) |
| CMS × Star-rating | 0.037 | -0.032 | -0.013 | -0.008 | -0.102 |
| | (0.067) | (0.064) | (0.075) | (0.070) | (0.080) |
| NHAA × Star-rating | 0.008 | 0.004 | -0.062 | -0.025 | -0.109 |
| | (0.068) | (0.065) | (0.076) | (0.071) | (0.081) |
| Constant | -1.678*** | -2.010*** | 0.270 | -1.490*** | 0.358* |
| | (0.153) | (0.145) | (0.170) | (0.158) | (0.181) |
| R-squared | 0.228 | 0.298 | 0.019 | 0.150 | 0.330 |
| *N* | 1578 | 1558 | 1601 | 1600 | 1607 |

Note: OLS regression. Standard errors are shown in parentheses. The reference group for CMS and NHAA is Google Health. Two-tailed tests, * p<0.05, ** p<0.01, *** p<0.001

**Table 5. Sector Bias and Star Ratings**

| | Efficiency | Effective | Red tape | Equity | Comfortable |
|---|---|---|---|---|---|
| Government | 0.039 | 0.183 | -0.035 | 0.176 | 0.119 |
| | (0.204) | (0.195) | (0.229) | (0.213) | (0.243) |
| Nonprofit | 0.175 | 0.316 | 0.056 | 0.339 | 0.331 |
| | (0.208) | (0.200) | (0.233) | (0.216) | (0.247) |
| Star-rating | 0.587*** | 0.701*** | -0.136** | 0.506*** | 0.924*** |
| | (0.046) | (0.044) | (0.052) | (0.048) | (0.055) |
| Government × Star-rating | 0.009 | -0.041 | 0.005 | -0.024 | -0.008 |
| | (0.066) | (0.063) | (0.074) | (0.069) | (0.078) |
| Nonprofit × Star-rating | -0.022 | -0.067 | -0.033 | -0.076 | -0.045 |
| | (0.067) | (0.064) | (0.075) | (0.069) | (0.079) |
| Constant | -1.820*** | -2.164*** | 0.430** | -1.590*** | 0.431* |
| | (0.144) | (0.137) | (0.161) | (0.149) | (0.171) |
| R-squared | 0.229 | 0.300 | 0.015 | 0.153 | 0.333 |
| *N* | 1,578 | 1,558 | 1,601 | 1,600 | 1,607 |

Note: OLS regression. Standard errors are shown in parentheses. The reference group for public and nonprofit nursing homes is for-profit homes. Two-tailed tests, * p<0.05, ** p<0.01, *** p<0.001

**Table 6: Bias in Star Ratings**

| | Efficiency | Effective | Red tape | Equity | Comfortable |
|---|---|---|---|---|---|
| 4-star | 0.560*** | 0.650*** | -0.098 | 0.472*** | 0.864*** |
| | (0.054) | (0.052) | (0.061) | (0.056) | (0.065) |
| 2-star | -0.607*** | -0.685*** | 0.194** | -0.477*** | -0.953*** |
| | (0.054) | (0.052) | (0.061) | (0.056) | (0.065) |
| Constant | 0.014 | 0.009 | -0.032 | 0.001 | 3.328*** |
| | (0.038) | (0.037) | (0.043) | (0.040) | (0.046) |
| R-squared | 0.227 | 0.297 | 0.015 | 0.149 | 0.329 |
| *N* | 1,578 | 1,558 | 1,601 | 1,600 | 1,607 |

Note: OLS regression. Standard errors are shown in parentheses. The reference group for 4-star and 2-star is 3-star. Two-tailed tests, * p<0.05, ** p<0.01, *** p<0.001

**Table 7: Evaluator and Credibility**

| | Credibility |
|---|---|
| CMS | 0.105+ |
| | (0.063) |
| NHAA | 0.142* |
| | (0.063) |
| Constant | -0.083+ |
| | (0.044) |
| R-squared | 0.004 |
| *N* | 1,520 |

Note: OLS regression. Standard errors are shown in parentheses. The reference group for CMS and NHAA is Google Health. Two-tailed tests, + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

**Table 8. Star Ratings and Credibility**

| | Efficiency | Effective | Red tape | Equity | Comfortable |
|---|---|---|---|---|---|
| Government | 0.059 | 0.060 | -0.021 | 0.105 | 0.091 |
| | (0.054) | (0.051) | (0.062) | (0.056) | (0.065) |
| Nonprofit | 0.086 | 0.084 | -0.031 | 0.077 | 0.185** |
| | (0.054) | (0.051) | (0.063) | (0.056) | (0.066) |
| Star-rating | 0.539*** | 0.626*** | -0.144*** | 0.421*** | 0.871*** |
| | (0.028) | (0.026) | (0.032) | (0.029) | (0.033) |
| Credibility | -0.641*** | -0.779*** | 0.386*** | -0.752*** | -0.659*** |
| | (0.084) | (0.080) | (0.097) | (0.087) | (0.102) |
| Star-rating × Credibility | 0.250*** | 0.291*** | -0.126*** | 0.288*** | 0.234*** |
| | (0.027) | (0.026) | (0.031) | (0.028) | (0.033) |
| Constant | -1.711*** | -1.976*** | 0.465*** | -1.376*** | 0.558*** |
| | (0.091) | (0.086) | (0.105) | (0.095) | (0.110) |
| R-squared | 0.283 | 0.366 | 0.026 | 0.214 | 0.354 |
| *N* | 1,491 | 1,470 | 1,514 | 1,512 | 1,519 |

Note: OLS regression. Standard errors are shown in parentheses. The reference group for public and nonprofit nursing homes is for-profit homes. Two-tailed tests, * p<0.05, ** p<0.01, *** p<0.001

# References


Aaker, Jennifer L., Emily N. Garbinsky, and Kathleen D. Vohs. 2012. Cultivating Admiration in Brands: Warmth, Competence, and Landing in the "Golden Quadrant." *Journal of Consumer Psychology* 22, no. 2: 191-194.

Abt Associates Inc. 2014. *Nursing Home Compare Five-Star Quality Rating System: Year Five Report*. Centers for Medicare & Medicaid Services.

Amirkhanyan, Anna A., Stephen B. Holt, Austin M. McCrea, and Kenneth J. Meier. 2019. Managing Racial Diversity: Matching Internal Strategies with Environmental Needs. *Public Administration Review* 79, no. 1: 69-81.

Amirkhanyan, Anna A., Hyun Joon Kim, and Kristina T. Lambright. 2008. Does the Public Sector Outperform the Nonprofit and For-profit Sectors? Evidence from a National Panel Study on Nursing Home Quality and Access. *Journal of Policy Analysis and Management* 27, no. 2: 326-353.

Amirkhanyan, Anna A., Hyun Joon Kim, and Kristina T. Lambright. 2010. Do Relationships Matter? Comparing the Performance of Relational and Classic Contracts. *Public Performance & Management Review* 34, no. 2: 189-220.

Amirkhanyan, Anna A., Kenneth J. Meier, Laurence J. O'Toole Jr, Mueen A. Dakhwe, and Shawn Janzen. 2018. Management and Performance in US Nursing Homes. *Journal of Public Administration Research and Theory* 28, no. 1: 33-49.

Andersen, Lotte Bøgh, Andreas Boesen, and Lene Holm Pedersen. 2016. Performance in Public Organizations. *Public Administration Review* 76, no. 6: 852-862.

Andrews, Rhys, George A. Boyne, and Richard M. Walker. 2011. Dimensions of Publicness and Organizational Performance: A Review of the Evidence. *Journal of public administration research and theory* 21, no. suppl_3: i301-i319.

Baekgaard, Martin, and Søren Serritzlew. 2016. Interpreting Performance Information: Motivated Reasoning or Unbiased Comprehension. *Public Administration Review,* 76, no. 11: 73-82.

Battaglio Jr, R. Paul, Paolo Belardinelli, Nicola Bellé, and Paola Cantarelli. 2019. Behavioral Public Administration Ad Fontes. *Public Administration Review* 79, no. 3: 304-320.

Ben-Ner, Avner, Darla J. Hamann, and Ting Ren. 2019. Does Ownership Matter in the Selection of Service Providers? Evidence from Nursing Home Consumer Surveys. *Nonprofit and Voluntary Sector Quarterly*, 47, no. 6: 1271-1295.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis*, 20, no 3: 351-368.

Bok, Derek. 2001. *The Trouble with Government*. Cambridge MA: Harvard University Press.

Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2016. Amazon's Mechanical Turk: A New Source of Inexpensive, yet High-quality Data? *Perspectives on Psychological Science*, 6, no. 1: 3-5.

Button, Katherine S., John PA Ioannidis, Claire Mokrysz, Brian A. Nosek, Jonathan Flint, Emma SJ Robinson, and Marcus R. Munafò. 2013. Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience. *Nature Reviews Neuroscience* 14, no. 5: 365–376.

Charbonneau, Étienne, and Gregg G. Van Ryzin. 2015. Benchmarks and Citizen Judgments of Local Government Performance. *Public Management Review* 17, no. 2: 288-304.

Christensen, Tom, and Per Lægreid. 2005. Trust in Government: The Relative Importance of Service Satisfaction, Political Factors, and Demography. *Public Performance & Management Review* 28, no. 4: 487–511.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

———. 1967. *Inside Bureaucracy.* Boston: Little, Brown.

Drevs, Florian, Dieter K. Tscheulin, and Jörg Lindenmeier. 2014. Do Patient Perceptions Vary With Ownership Status? A Study of Nonprofit, For-Profit, and Public Hospital Patients. *Nonprofit and Voluntary Sector Quarterly*, 43 no. 1: 164–184.

Eggleston, Karen, and Richard Zeckhauser. 2002. Government Contracting for Health Care. In J. D. Donahue & J. S. Nye, Jr. (Eds.), *Market-based Governance* (pp. 29–65). Washington, DC: Brookings Institution Press.

Feeney, Mary K., and Hal G. Rainey, Hal G. 2010. Personnel Flexibility and Red Tape in Public and Nonprofit Organizations. *Journal of Public Administration Research and Theory*, 20, no. 4: 801-826

Fiorillo, V. (2015). Philly Woman Sues Angie's List, Calls Service "Fraudulent," "Deceptive" Claims the company secretly manipulates reviews and ratings. Philadelphia, accessed on April 8, 2020 on the web site: https://www.phillymag.com/news/2015/03/26/angies-list-class-action-lawsuit/

Frederickson, H. George. 1971. Toward a New Public Administration. In Toward a New Public Administration: The Minnowbrook Perspective, edited by Frank Marini, 309–31. Scranton, PA: Chandler.

Goodsell, Charles T. 2015. *The New Case for Bureaucracy: A Public Administration Polemic*. Thousand Oaks, CA: Sage.

Hall, Jeremy L., and Gregg G. Van Ryzin. 2019. A Norm of Evidence and Research in Decision-making (NERD): Scale Development, Reliability, and Validity. *Public Administration Review* 79, no. 3: 321-329.

Handy, Femida, Stephanie Seto, Amanda Wakaruk, Brianna Mersey, Ana Mejia, and Laura Copeland. 2010. The Discerning Consumer: Is Nonprofit Status a Factor? *Nonprofit and Voluntary Sector Quarterly*, 39 no. 5: 866–883.

Hansman, Henry. 1996. The Changing Roles of Public, Private, and Nonprofit Enterprise in Education, Health Care, and Other Human Services. In V. R. Fuchs (Ed.), *Individual and Social Responsibility: Child Care, Education, Medical Care, and Long-term Care in America* (pp. 245–275). Chicago: The University of Chicago Press.

Harrington, Charlene, Steffie Woolhandler, Joseph Mullan, Helen Carrillo, and David U. Himmelstein. 2001. Does Investor Ownership of Nursing Homes Compromise the Quality of Care? *American Journal of Public Health*, 91 no. 9: 1452–55.

Health Deficiencies. n.d. Data. Medicare.gov. Accessed on April 8, 2020 on the web site: https://data.medicare.gov/Nursing-Home-Compare/Health-Deficiencies/r5ix-sfxw

Hodge, Graeme. 2018. *Privatization: An International Review of Performance*. New York: Routledge.

Holm, Jakob Majlund. 2017. Double standards? How Historical And Political Aspiration Levels Guide Managerial Performance Information Use. *Public Administration*, 95 no. 4: 1026-1040.

Hong, Sounman, and Youngrok Kim. 2019. Loyalty or Competence: Political Use of Performance Information and Negativity Bias. *Public Administration Review* 79, no. 6: 829-840.

Hood, Christopher, and David Heald. 2006. *Transparency: The Key to Better Governance?* Oxford, UK: Oxford University Press.

Hvidman, Ulrik. 2019. Citizens' Evaluations of the Public Sector: Evidence From Two Large-Scale Experiments. *Journal of Public Administration Research and Theory,* 29 no. 2: 255–267.

Hvidman, Ulrik and Simon Calmar Andersen. 2016. Perceptions of Public and Private Performance: Evidence from a Survey Experiment. *Public Administration Review*, 76 no. 1: 111–120.

James, Oliver, Sebastian Jilke, Carolyn Petersen, and Steven Van de Walle. 2016. Citizens' Blame of Politicians for Public Service Failure. *Public Administration Review*, 76 no. 1: 83-93.

James, Oliver, and Gregg G. Van Ryzin. 2016. Motivated Reasoning about Public Performance. *Journal of Public Administration Research and Theory*, 27 no. 1: 197-209.

James, Oliver, and Gregg G. Van Ryzin. 2017. Incredibly Good Performance: An Experimental Study of Source and Level Effects on the Credibility of Government. *The American Review of Public Administration*, 47 no. 1: 23-35.

Jilke, Sebastian, Wouter Van Dooren, and Sabine Rys. 2018. "Discrimination and Administrative Burden in Public Service Markets. *Journal of Public Administration Research and Theory* 28, no. 3: 423-439.

Jilke, Sebastian, Gregg G. Van Ryzin, and Steven Van de Walle. 2016. Responses to Decline in Marketized Public Services. *Journal of Public Administration Research and Theory*, 26 no. 3: 421–432.

Keiser, Lael R., and Susan M. Miller. 2020. Does Administrative Burden Influence Public Support for Government Programs? *Public Administration Review* 80, no. 1: 137-150.

Marvel, John D. 2015. Public Opinion and Public Sector Performance: Are Individuals' Beliefs About Performance Evidence-Based or the Product of Anti–Public Sector Bias? *International Public Management Journal*, 18, no. 2: 209–227.

Marvel, John D. 2016. Unconscious Bias in Citizens' Evaluations of Public Sector Performance. *Journal of Public Administration Research and Theory*, 26, no. 1: 143–158.

Marvel, John D. and Amanda Girth. 2016. Citizen Attributions of Blame in Third-party Governance. *Public Administration Review*, 76, no. 1: 96-108.

Meier, Kenneth J., Austin Johnson, and Seung-ho An. 2019. Perceptual Bias and Public Programs: The Case of the United States and Hospital Care. *Public Administration Review*, 79 no. 6: 820-828.

Moynihan, Donald P. 2006. What Do We Talk about When We Talk about Performance? Dialogue Theory and Performance Budgeting. *Journal of Public Administration Research and Theory*, 16, no. 2: 151–68.

Nagtegaal, Rosanna, Lars Tummers, Mirko Noordegraaf, and Victor Bekkers. 2020. Designing to Debias: Measuring and Reducing Public Managers' Anchoring Bias. *Public Administration Review* 80, no. 4: 565-576.

National Center for Health Statistics. 2019. Long-term Care Providers and Services Users in the United States, 2015–2016. Retrieved from https://www.cdc.gov/nchs/data/series/sr_03/sr03_43-508.pdf

Nielsen, Poul A., and Donald P. Moynihan. 2017. How Do Politicians Attribute Bureaucratic Responsibility for Performance? *Journal of Public Administration Research and Theory*, 27 no. 2: 269–83.

Niskanen, William A. 1971. *Bureaucracy and Representative Government*. Chicago: Aldine-Altherton.

Olsen, Asmus L. 2017a. Human Interest or Hard Numbers? Experiments on Citizens' Selection, Exposure, and Recall of Performance Information. *Public Administration Review*, 77, no. 3: 408-420.

Olsen, Asmus Leth. 2017b. Compared to What? How Social and Historical Reference Points Affect Citizens' Performance Evaluations, *Journal of Public Administration Research and Theory*, 27, no 4: 562-580.

O'Neill, Ciaran, Charlene Harrington, Martin Kitchener, and Debra Saliba. 2003. Quality of Care in Nursing Homes. *Medical Care* 41, no. 12: 1315–22.

Paolacci, Gabriele, and Jesse Chandler. 2014. Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*, 23, No. 3: 184–188.

Persson, Torsten, and Guido Tabellini. 1999. The Size and Scope of Government: Comparative Politics with Rational Politicians. *European Economic Review*, 43, no. 4-6: 699-735.

Pew Research Center. (2019). Public Trust in Government: 1958-2019 Retrieved from https://www.people-press.org/2019/04/11/public-trust-in-government-1958-2019/

Pornpitakpan, Chanthika. 2004. The Persuasiveness of Source Credibility. *Journal of Applied Social Psychology*, 34, no. 2: 243-281.

Riccucci, Norma M. 2009. The Pursuit of Social Equity in the Federal Government: A Road Less Traveled? *Public Administration Review*, 69 no. 3: 373-382.

Riportella-Muller, Roberta, and Doris P. Slesinger. 1982. The Relationship of Ownership and Size to Quality of Care in Wisconsin Nursing Homes. *The Gerontologist* 22, no. 4:429–34.

Rölle, Daniel. 2017. Mass Media and Bureaucracy-bashing: Does the Media Influence Public Attitudes towards Public Administration? *Public Policy and Administration*, 32, no. 3, 232-258.

Rothbart, Michah W., Amy Ellen Schwartz, and others. 2019. What a Difference a Grade Makes: Evidence from New York City's Restaurant Grading Policy. *Public Administration Review* 79, no. 5: 651-665.

Santerre, Rexford E., and John Vernon. 2005. Testing for Ownership Mix Efficiency: The Case of the Nursing Home Industry. Working Paper no. 11115, National Bureau of Economic Research. http://www.nber.org/papers/w11115.

Savas, E.S. 2000. *Privatization and Public-private Partnerships*. New York: Chatham House Publishers.

Schlesinger, Mark and Bradford H. Gray. 2006. Nonprofit Organizations and Health Care: Some Paradoxes of Persistent Scrutiny. In The *Nonprofit Sector: A Research Handbook*, 2nd ed., edited by Walter W. Powell and Richard Steinberg, 378-414. New Haven, CT: Yale University Press.

Schlesinger, Mark, Shannon Mitchell, and Bradford H. Gray. 2004. Restoring Public Legitimacy to the Nonprofit Sector. *Nonprofit and Voluntary Sector Quarterly*, 33 no. 4:673–710.

Stritch, Justin M., Mogens Jin Pedersen, and Gabel Taggart. 2017. The Opportunities and Limitations of Using Mechanical Turk (Mturk) in Public Administration and Management Scholarship. *International Public Management Journal*, 20 no. 3: 489–511.

Van Slyke, David M., and Christine H. Roch. 2004. What Do They Know, and Whom Do They Hold Accountable? *Journal of Public Administration Research and Theory*, 14 no.2: 191-209.

Van Ryzin, Gregg G., and Cecilia F. Lavena. 2013. The Credibility of Government Performance Reporting. *Public Performance & Management Review*, 37 no. 1: 87-103.

Yang, Kaifeng, and Marc Holzer. 2006. The Performance–trust Link: Implications for Performance Measurement. *Public Administration Review*, 66 no. 1:114-126.

Yoon, Kak, Choong Hyun Kim, and Min-Sun Kim. 2011. A Cross-cultural Comparison of the Effects of Source Credibility on Attitudes and Behavioral Intentions. *Mass Communication and Society*, 1, no. 3-4: 153-173.

Walker, Richard M., Gene A. Brewer, M. Jin Lee, Nicolai Petrovsky, and Arjen Van Witteloostuijn. 2019. Best Practice Recommendations for Replicating Experiments in Public Administration. *Journal of Public Administration Research and Theory*, 29 no. 4: 609-626.

Watson, Sidney D. 2009. From Almshouses to Nursing Homes and Community Care: Lessons from Medicaid's History. *Georgia State University Law Review*, 26 no. 3: 937–970.

Willems, Jurgen. 2020. Public Servant Stereotypes: It Is Not (at) all About Being Lazy, Greedy and Corrupt. *Public Administration* 98, no. 4:807-823.

Zhang, Jiasheng, Wenna Chen, Nicolai Petrovsky, and Richard M. Walker. 2021. The Expectancy-Disconfirmation Model and Citizen Satisfaction with Public Services: A Meta-analysis and an Agenda for Best Practice. *Public Administration Review* doi 10.1111/puar.13368.