You Can't Always Get What You Want:

Data Access in US Small and Medium Sized Cities

by

Federica Fusi

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Public Administration

Approved September 2018 by the
Graduate Supervisory Committee:

Mary K. Feeney, Chair
Stephan Grimmelikhuijsen
Karen Mossberger
Eric W. Welch

ARIZONA STATE UNIVERSITY

December 2018

ABSTRACT

This research examines data exchange between city departments and external stakeholders; particularly, why city departments have different capacity to access data from departments in the same city, other public agencies, private and nonprofit organizations. Data access is of theoretical interest because it provides the opportunity to investigate how public organizations and public managers deal with a portfolio of relationships in a loosely structured context characterized by dynamics of power and influence. Moreover, enhancing data access is important for public managers to increase the amount and diversity of information available to design, implement, and support public services and policies.

Drawing from institutionalism, resource dependence theory, and collaboration scholarship, I developed a set of hypotheses that emphasize two dimensions of data access in local governments. First, a vertical dimension which includes institutions, the social environment - particularly power relationships - and coordination mechanisms implemented by managers. This dimension shows how exogenous - not controlled by public managers - and endogenous - controlled by public managers - factors contribute to a public organization's ability to access resources. Second, a horizontal dimension which considers the characteristics of the actors involved in data exchange and emphasizes the institutional and social context of intra-organizational, intra-sectoral and cross-sectoral data access.

Hypotheses are tested using survey data from a 2016 nationally representative sample of 500 US cities with populations between 25,000 and 250,000. By focusing on

small- and medium-sized cities, I contribute to a literature that typically focuses on data sharing in US large cities and federal agencies. Results show that the influence of government agencies and the influence of civil society have opposite effect on data access, whereas government influence limits data access while influence from civil society increases capacity to access data. The effectiveness of coordination mechanisms varies according to the stakeholder type. Public managers rely on informal networks to exchange data with other departments in the city and other governmental agencies while they leverage lateral coordination mechanisms - informal but unplanned - to coordinate data access from nongovernmental organizations. I conclude by discussing the implications for practice and future research.

ACKNOWLEDGMENTS

The data for this research project were collected in 2016 and provided by the Center for Science, Technology, and Environmental Policy Studies (CSTEPS) at Arizona State University and The Pew Charitable Trusts. The view expressed herein are those of the author and do not necessarily reflect the views of CSTEPS and The Pew Charitable Trusts.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Researchers and public managers have shown growing interest in government data practices in the past decade. Examples of data practices include Open Data web portals implemented by the US federal government[1] and major US cities,[2] as well as initiatives to develop new methods for data collection, such as the Big Data working group promoted by the US Census Bureau[3]. Data practices also include partnerships between government agencies and external stakeholders, such as the one initiated by the California Office to Reform Education (CORE) districts to collect and share data among California major school systems[4]. Many other initiatives oriented towards "increase the amount or quality of data collected, better share them across or use them within agencies, or improve capacity to analyze and distribute summary statistics" can be found at local, state, and national level (Weitzman, Silver and Brazil, 2006, p. 387).

City governments widely support data practices because they expect that improvements in data collection, diffusion, and use will lead to significant benefits for city activities and design and delivery of public services (Gil-Garcia & Sayogo, 2016). For instance, police departments access data from national security agencies to coordinate

---

[1] The Open Data portal of the US government is available here: https://www.data.gov/

[2] For example, see the city of Phoenix - https://www.phoenix.gov/opendata or the city of Los Angeles - https://data.lacity.org/

[3] For more information: https://www.census.gov/about/cac/sac/wg-big-data.html

[4] The CORE initiative is described here: http://coredistricts.org/

local interventions; finance departments use employment and hiring data from other departments to manage budgets and pay employees; community development departments rely on data about businesses, nonprofit organizations, schools, and other local initiatives to allocate resources and promote citizen involvement; and so on. Improving data collection and access can significantly reduce costs and increase the information available (Dawes, 1996).

Moreover, city governments have gained a central role in the current policy environment characterized by high cross-organizational interdependence, rapid changes, participation, and transparency pressures. Cities have expressed great interest in data practices to improve their relationships with stakeholders (Kitchin, 2014). For instance, city managers leverage data to incentivize bottom-up processes of co-production and increase interaction with civil society actors (Attard, Orlandi, Scerri, & Auer, 2015). Open data portals aim at engaging citizens and promoting the development of citizen-driven innovation and proposals for public policies. Open data portals also enhance interactions among public and private actors, from merging complementary datasets to using private data for better understanding matters of public interest (Susha, Janssen, & Verhulst, 2017). Data practices related to openness are oriented towards facilitating government data access and use and improving the quality and participative nature of decision-making.

Data practices are a substantial research area for public management scholars. Several studies look at how cities are managing their data (Kitchin, 2014; Roberts, 2011; Susha et al., 2017), which challenges they face in exchanging data across stakeholders

(Dawes, 1996; Gil-Garcia et al., 2007), and how they use data for public purposes (Attard et al. 2015; Chan, 2013; Clarke & Margetts, 2014). This scholarship recognizes that data management faces different challenges and requires different solutions in public organizations than in private ones, as public managers pay greater attention to issues related to transparency, accountability, privacy, and security (Bozeman & Bretschneider, 1986; Clarke & Margetts, 2014; Meijer, 2015). Nevertheless, research in this area is relatively new, and several theoretical and empirical gaps call for further research to understand data practices in different organizational contexts and for different purposes; disentangle theoretical implications; and provide practical advice to government.

This research finds its niche by investigating data access across small- and medium-sized cities in the US. Evidence from a survey conducted in 2016 on a nationally representative sample of 500 cities with populations between 25,000 and 250,000, shows that city departments cannot timely access data from other organizations and that data access decreases when data requests are submitted to non-governmental organizations (Feeney et al., 2016). Investigating such variation is essential to understand how to enhance city capacity to collect data and improve the availability of information for public managers and local politicians. I ask: Why are some city departments more able to access data from other organizations? What role do the institutions, the social environment, and coordination mechanisms play in shaping data access? To answer these questions, I develop an Integrative Framework for Data Access (IFDA) that explains the main factors shaping data access in the public sector.

The IFDA combines a vertical and a horizontal dimension. The vertical dimension includes the city's institutional and social environment as well as coordination mechanisms used by local public managers. I suggest that while institutions and the social environment shape constraints and opportunities to access data at the macro-level, at the micro-level city managers design agreements and communicate with other managers to access data from other organizations. Despite their relevance in inter-organizational research, few studies have investigated coordination in data sharing (Nedović-Budić & Pinto, 2000; Susha et al., 2017). This approach complements current literature, which mostly examines the implementation and use of technology tools and infrastructures while overlooking how formal and informal mechanisms might coexist with technology to facilitate access to data.

The horizontal dimension highlights differences across relationship types. City departments can request data from departments within their city, other public agencies, and non-governmental organizations. Accessing data from the full portfolio of relationships increases data available to local managers and policymakers. Yet each relationship type is characterized by different social and institutional aspects that might hinder or facilitate data access. The horizontal dimension expands public management knowledge on how cities can access data from other organizations and how coordination mechanisms vary according to the type of relationship.

I describe in more detail the context, research questions, theoretical approach, and contributions adopted in this study in the remaining of the introduction. First, I describe the technological, social and organizational context of data practices and how data practices have become relevant for public management researchers and practitioners, specifically in city governments. Then, I discuss the research questions and the theoretical and **e**mpirical motivation of the study. Finally, I briefly illustrate the theoretical framework and discuss contributions for theory and practice.

Research Context

The growing diffusion of data practices among government agencies is the response to three phenomena that have strongly influenced government activities in recent decades: the advancement of information and communication technologies (ICTs) which offer new tools for data management at a lower cost; the growing complexity of policy problems requiring several organizations to coordinate and share information on their activities; and citizens' demands for transparency and accountability of governments, which lead politicians and public managers to seek for better ways to diffuse information. Synergies and interconnections among these phenomena have increased attention to data practices.

The growth of information and communication technologies (ICTs) has provided new tools for the collection, storage, sharing, and use of data. For instance, governments can now collect data more efficiently and on a larger population via online surveys. They can also leverage new sources of data, such as traffic and mobile data, which provide more granular and frequently updated information and at lower costs (Bostic, 2015). Cloud-based platforms such as Dropbox and Google Drive have created the opportunity to simultaneous share data with multiple actors – including citizens - who can access the same digital archive from their own electronic devices. Finally, new techniques based on machine learning and predictive analytics have simplified the analysis of real-time, high volume data to inform policy and decision- making processes. All in all, these developments have significantly reduced the costs and the complexity of collecting, using and sharing data, and public agencies are exploring ways to take advantage of these opportunities.

Public managers expect that a greater availability of data will help them better understanding complex and rapidly changing policy problems, which government agencies struggle to address on their own. Rittel and Webber (1973) coined the term "wicked problems" to refer to societal problems – such as poverty, crime or social inclusion – whose definitions are highly ambiguous and socially constructed. Wicked problems require the involvement of several actors – across government levels, sectors and national boundaries - to guarantee legitimacy and implementation of policy solutions (Ansell & Gash, 2008; Jasanoff & Martello, 2004; Koppenjan & Klijn, 2004). Sharing data across organizations is a first critical step to integrate activities and processes and achieve common goals (Dawes, 1996; Gil-Garcia & Sayogo, 2016; Guo & Acar, 2005; Roberts, 2011). Growing urbanization, for instance, has created large metropolitan areas which require cities to coordinate and share data to provide public services, including transportation, welfare assistance, and public safety (Lefèvre, 1998; Parks & Oakerson, 1989). Environmental policies are also implemented thanks to the combined action of multiple federal, state and local governments and nongovernmental actors who collaboratively manage and protect natural resources (Lubell, Schneider, Scholz, & Mete, 2002).

Along with the complexity of policy problems, the specialization of tasks among policy actors has expanded the need to collaborate across disciplines and sectors (Hale, 2011; Koppenjan & Klijn, 2004). Emergency policies, for instance, require integration of knowledge and expertise from citizens, non-governmental organizations (NGOs), state and local agencies to coordinate short and long-term interventions in the affected areas (Roberts, 2011). Cross-discipline collaboration is also typical in public health, where homeland security departments collaborate with public health and governmental institutions to prevent biosafety and bioterrorism concerns (Daley, 2009). Governments have incentives to share data to reduce the costs associated with data collection and analysis, avoid duplication of efforts, build policy solutions based on integrated expertise and knowledge, and gain insights on policy issues (Dawes, 1996). Although government officials are aware that socio-political factors might affect policy outcomes, they often perceive data as a final input to decision making (Weitzman, Silver, & Brazill, 2006).

Finally, data practices are crucial to addressing societal pressures coming from citizens, nonprofit organizations, and other external stakeholders. Governments face rising mistrust from citizens (Pew Research, 2017) and leverage open data initiatives to increase transparency and accountability and regain citizen confidence (Bertot, Jaeger, & Grimes, 2010). Similarly, governments seek to expand participation in decision making and policy design by sharing information with the public (Grimmelikhuijsen & Welch, 2012; Hood, 2006; Meijer, 2015). The democratic model of government is slowly turning away from a representation-based approach, where citizens are called to elect and monitor their representatives, towards a participation-oriented approach, where citizens are directly involved in government activities (Meijer, 2015). Researchers suggest that governments should increase open data to support bottom-up participation and citizen-driven innovation processes (Attard et al., 2015).

In summary, public managers have emphasized data practices because government agencies are facing a policy environment where actors are more interconnected to one another, and policy problems are more nuanced and loosely defined. Technology advancements promise the tools to digitally represent and interpret this increasingly complex reality by easing the collection, storage, and analysis of data. Information technologies offer “a latent capacity to share information across agency and program boundaries, to discover patterns and interactions once hidden in millions of separate paper records and to make decisions based on more complete data” (Dawes, 1996, p. 377). Moreover, new IT tools provide opportunities to shift from top-down decision-making processes to bottom-up processes of co-production and interaction with civil society actors, therefore responding to pressures for accountability and transparency (Bertot et al., 2010; Desouza & Bhagwatwar, 2014).

Data practices in city government. City governments find themselves at the center of rapid ICT developments, high cross-organizational interdependence, and pressure to increase participation and transparency. Consequently, they have expressed great interest in data practices to improve public service provision, manage their relationships with stakeholders, and address citizen concerns (Desouza & Bhagwatwar, 2014; Kitchin, 2014; Meijer & Thaens, 2018).

With the global population becoming predominantly urban (UN Habitat, 2011), city governments emerge as the primary providers of public services. City governments have broad authority over critical public issues such as infrastructure development, energy, environment, migration, and transportation. They also have a vast influence on citizens' life and behavior as their status of “local, directly elected bodies” put them in a unique leadership position (Department of the Environment, Transport and the Regions, 2000, p. 40). Moreover, global policies often lack authority to impact individuals directly and rely on a combination of local and global actions to achieve their goals. For instance, city governments are key actors in climate change policies because they have authority to enforce transportation practices and rules to reduce emission levels (Betsill & Bulkeley, 2004, 2006; Pitt & Bassett, 2013). Recently, some city mayors have taken a strong stance against the US federal government’s decision to withdraw from the Paris agreement on climate change (Walker, 2017) by renewing their commitment to its goals.

City managers and politicians have warmly welcomed improvements in data collection and analysis, hoping a greater availability of data will aid to respond to citizen’s needs and global pressures (Townsend, 2013). City managers assume that more up-to-date, timeless, and granular data will lead to new and innovative solutions to manage wicked problems and interdependences (Bettencourt, 2014; Dodgson and Gann, 2011). So-called “smart cities” are an example of such intense use of data practices. Smart cities leverage in situ and remote sensors to collect real-time information on urban phenomena and inform policy design (Koonin & Holland, 2014). Cities are also asking citizens to provide information about urban problems and analyze data (Desouza &

Bhagwatwar, 2014; Kim & Lee, 2012). Citizens can use mobile applications (e.g., "apps") to share information about their neighborhood (Ertiö, 2015) or they can download data from open data portals and propose local innovations and policies (Attard et al., 2015; Chan, 2013).

City governments, however, face difficulties in implementing and adopting data practices. Practitioners and researchers highlight how managerial, organizational, institutional, and technical barriers can hold back city government efforts (Gil-Garcia & Sayogo, 2016; Grimmelikhuijsen & Feeney, 2016; Roberts, 2011; Tulloch & Harvey, 2007; Welch, Feeney, & Park, 2016; Yang & Wu, 2016). In particular, small city governments lack financial and technical resources and managerial capacity to support initiatives such as smart cities or data sharing partnerships (Gil-Garcia & Sayogo, 2016; Homburg & Bekkers, 2002; Welch et al., 2016). Furthermore, they might lack support from local politicians scared of exposing government data because of privacy issues or concerns with negative criticism (Attard et al., 2015; Janssen, Charalabidis, & Zuiderwijk, 2012). For these reasons, city managers might refrain from participating in data sharing initiatives, expect lower benefits from data use and exchange, and discourage data practices (de Montalvo, 2003; Gil-Garcia et al., 2007).

Research on data sharing should aim to understand how and what factors help city governments to overcome existing barriers and encourage data practices. City governments play a key role in providing data and information to the state and federal government. By working closely with local actors, city governments have access to local-level data that are too costly to be collected by other public agencies. Increasing data access for city governments is expected to have positive repercussions on the whole public sector. Moreover, data available to local governments directly impact decisions that affect citizens' everyday life and contribute to improve the quality of life of US residents.

Research Questions and Motivation

This research investigates a specific aspect of data practices, namely data access. Data access is an organization's capacity to timely obtain data from other organizations (further explanation: Chapter 3). Data access is essential because government agencies are often unable to autonomously collect the data they need, either because data collection is too expensive, or other organizations control the data. To overcome these problems, public managers request data from other organizations (Levine & White, 1961). For instance, city departments might request data from local nonprofits on clients served or at-risk populations. Data access increases the diversity and availability of data, which in turn enhance government effectiveness, efficiency, and responsiveness to local needs (Rothenberg & Zyglidopoulos, 2007). However, there is variation in how successful city departments are in obtaining data from other organizations (Nedović-Budić & Pinto, 2000). The following two anecdotes illustrate how data access can vary across public agencies.

Some years ago, I worked for an Italian local government where the community development department was legally required to regularly collect data from local NGOs (e.g., the number of members or the number of awarded grants). To my surprise, nonprofit organizations sent their data via mail or brought them in person to the department officers. The government lacked an online shared infrastructure, which significantly increased the time and resources required to collect the data. As a result, the data were often not up-to-date, and the department personnel spent considerable hours to contact nonprofit managers and remind them to send the data. While the department was ultimately able to access the data, employees faced inefficiencies, high workloads, and costs.

More recently, I spoke with a colleague who worked for a major nonprofit in Arizona. She coordinated a data sharing system between local nonprofit organizations and the city government. In stark contrast with the Italian government, local nonprofits in Arizona rely on a cloud-based platform to exchange and access data. Each organization can upload its data as well as access and download data posted by others. Participants consider the system to be efficient because it simplifies data collection and facilitates the coordination of joint activities. They also note that the system requires trust and competence among users and assumes that the data are not subject to privacy and security concerns that could make open sharing undesirable or illegal.

These anecdotes illustrate how some public organizations easily access data from external organizations while some others have irregular and fragmented access to such data, resulting in wasted time, stressed human resources, and outdated information. Similar findings emerge from previous research (Allard et al., 2018; Feeney et al., 2016; Pew Charitable Trusts, 2018), suggesting that data access varies across public agencies. Allard and colleagues (2018) found that some public managers report that they can generally access data they need for their work while others complain several delays. A study of a national sample of 500 small and medium-sized US cities confirms that most, but not all, city departments can timely access data from other organizations, but barriers increase when city governments request data from non-governmental organizations (Feeney et al., 2016). Based on this empirical evidence, this research focuses on data access to investigate the following research questions:

1. Why are some city departments more able to access data from other organizations?
2. What role do the institutions, social environment, and coordination mechanisms play in shaping data access?

An Integrative Theoretical Framework

The study proposes an integrative theoretical framework to explain data access in the public sector. The framework is entitled "Integrative Framework for Data Access" (IFDA) and includes two dimensions. The vertical dimension draws from previous research on institutions, resource dependency theory, collaboration studies, and information sharing. It argues that data access is a function of three dimensions. At the highest level, we find institutions that shape constraints and opportunities to access and exchange data across organizations. Then, the social environment shapes relationships across actors involved in the data exchange; in particular, influence dynamics might prevent organizations from providing data and therefore decrease access to data. Finally, at the micro-level, managers implement coordination strategies to access data. Strategies include formal and informal coordination mechanisms – from routines to personal relationships with other managers. Qualitative studies often mention these strategies, but we lack an empirical and theoretical understanding of how they structure data flows.

The horizontal dimension expands the framework across the full portfolio of stakeholders from which the city government might request data. While accessing data from a variety of organizations increases data availability and diversity, public managers face more challenges and barriers that might hinder data access. For instance, previous studies have shown that exchanging data across sectors is often complicated by cultural and value differences (Dawes et al., 2009; Roberts, 2011). Public management scholarship needs to expand our understanding of how social and institutional characteristics of cross-sector or intra-organizational relationships might affect data access. I consider three types of relationships: data access from other departments in the same city; data access from other public agencies; and data access from non-governmental organizations. The horizontal dimension intersects with the vertical one when I discuss the coordination mechanisms for data access; public managers will report greater (or lower) data access depending on which organization they are asking data from and the coordination mechanisms they adopt.

Contributions for Theory and Practice

This research contributes to public management and data sharing scholarship both theoretically and empirically. Theoretically, it focuses on data access, which public management research has rarely investigated. Data access provides insights into how public agencies can manage relationships with autonomous external stakeholders in a context where they might lack incentives to collaborate. The IFDA emphasizes elements at three different levels - institutions, social environment, and coordination mechanisms - to show how exogenous – not controlled by public managers – and endogenous – controlled by public managers – factors contribute to an organization's ability to access resources. Empirically, it provides implications for city managers that wish to improve their external relationships, and for state and federal initiatives that support and promote data sharing initiatives. Finally, few studies apply quantitative methodologies to investigate data practices in small and medium-sized cities.

Contribution to public management theory and research. Several studies have looked at data sharing initiatives, but few have investigated broad issues of data accessibility in the public sector. Gil-Garcia, Pardo, and Burke (2010) classify data sharing initiatives according to their focus. Some initiatives have a narrow focus, either they target a need (e.g., facilitate data access among a group of public agencies) or a policy problem (e.g., terrorist threats). Other initiatives have a broad focus and build systemic capacity for public agencies to access data for their daily activities. They found that the design of the initiative (broad vs. narrow focus) shapes the factors that matter for success. For instance, initiatives with a broad focus need trusted networks of actors while initiatives with a narrow focus rely on the deployment of effective infrastructure and integrated data. Up to now, research has mostly focused on narrowly focused initiatives, and we need more studies that investigate how to build systemic data access capacity in public agencies. By addressing this gap, this research aims to help public organizations to increase data availability and diversity.

Second, data access provides an opportunity to investigate how public organizations deal with their portfolio of stakeholders when relationships are loosely structured and likely characterized by power and influence dynamics (Willem & Buelens, 2007). Stakeholders are not required to share data with the government and might lack incentives for accepting data requests (Azad & Wiggins, 1995; Willem & Buelens, 2007). This research focuses on how institutions, social environment, and coordination mechanisms influence the likelihood to access data when public managers interact with different types of stakeholders. Studies on collaboration between public agencies and external stakeholders suggest that willingness to collaborate and successful outcomes depend on the typology of actors involved (e.g., private versus public actors or hierarchical vs. peer-to-peer relationships), but research on data sharing rarely distinguishes between the types of actors involved (Guo & Acar, 2005; Mullin & Daley, 2010; Roberts, 2011).

Third, this research emphasizes the role of managers and how they coordinate data access across the portfolio of stakeholders. The examples discussed above show that city managers adopt different coordination mechanisms to exchange data with other organizations, leading to different outcomes concerning efficiency, frequency, and effectiveness of the exchange. Several contributions provide a theoretical foundation to understand how coordination mechanisms enhance data sharing (Dawes, 1996; Gil-Garcia & Sayogo, 2016; Kim & Lee, 2006; Tulloch & Harvey, 2007; Welch et al., 2016; Willem & Buelens, 2007; Yang & Maxwell, 2011). For instance, both Gil-Garcia and colleagues (2010) and Yang and Maxwell (2011) highlight how personal networks are critical to explaining data and information sharing across organizations. Welch et al. (2016) suggest that coercion and persuasion mechanisms shape data sharing relationships. I further expand this literature by suggesting that the effectiveness of these coordination mechanisms might differ across stakeholders.

Contribution to practices. This research provides insights for practitioners and public managers. Results inform public managers about how to effectively manage data access with different types of organizations – namely, other departments in the same city, other public agencies, and non-governmental organizations, either nonprofit or for profit. Since the coordination mechanisms investigated in this study require limited resources, findings are implementable by department heads and city managers.

Moreover, national and state agencies are subsidizing infrastructures and programs to promote data and information sharing at the local level. Several foundations are also providing incentives and supporting initiatives on data practices (e.g., the What Works Cities initiative by the Bloomberg Foundation[5]). Without a systematic understanding of the factors that positively or negatively affect data access, it may be hard to direct public and nonprofit organizations funding; identify local governments that require greater support because of the challenges they are facing; and provide policy suggestions on how cities might intervene at the micro level to reduce barriers for cooperation and exchange. Results from this study provide research-based evidence that can help orient interventions of public and nonprofit organizations at a large scale and identify city governments that need additional resources.

[5] For more information, see: https://whatworkscities.bloomberg.org/

Finally, most research on data sharing has focused on national or state government agencies (Dawes et al., 2009; Dawes, Gharawi, & Burke, 2012; Higgins, Taylor, Lisboa, & Arshad, 2014) and has adopted a case study approach (Gil-Garcia et al., 2007; Roberts, 2011). Few studies have collected quantitative data on data access among cities in general, and specifically small- and medium-sized cities (exceptions include Gil-Garcia & Sayogo, 2016 and Welch et al., 2016, among others). Small and medium-size cities have lower financial resources and technical capacity (Gil-Garcia & Sayogo, 2016; Hamin, Gurran, & Emlinger, 2014; Homburg & Bekkers, 2002) so as they are less likely to be included in large data sharing initiatives, lead them, or build data sharing infrastructure (Nedović-Budić & Pinto, 2000). This research offers insights for public managers and practitioners interested in building capacity for data exchange and access within small and medium-sized cities.

Dissertation Structure

The dissertation is organized as follows. Chapter 2 presents current research on data sharing and access. First, it describes data, information, and knowledge sharing studies in organizational studies and public management scholarship. Then, it defines "data" and discusses data used in public organizations. Finally, it presents theoretical and empirical findings from previous research; it concludes by highlighting the main research gaps addressed in this study.

Chapter 3 presents the IFDA. It introduces the theoretical rationales behind the vertical and horizontal dimensions and frames them within current research in public management. It discusses the importance of applying the IFDA to investigate data access in small and medium-sized US cities.

Chapter 4 presents the research hypotheses. The hypotheses discuss institutions (privacy laws, legal mandate, open data quality, and institutional capacity); social environment (external and internal influence); coordination mechanisms (formal, lateral, informal coordination and technology tools); and the intersection between coordination mechanisms and stakeholder type (other departments in the same city, other public agencies, and non-governmental organizations).

Chapter 5 presents the data collection and variable measurement. Data were collected as part of a 2016 national survey conducted on 500 small- and medium-sized US cities by the Center for Science, Technology, and Environmental Policy Studies at Arizona State University. The chapter describes the sample frame, survey design, data collection, and measurement of the variables. The chapter also introduces state-level data that were collected from external organizations and employed in the research. The chapter reviews the methodology utilized for the data collection and describes the variables.

Chapter 6 discusses data analysis and presents the results of the empirical models. Three different statistical methods were used to analyze the data: a logit model with clustered robust standard errors; a multi-level logit model; and a seemingly unrelated regressions model. The chapter compares the three models and discusses their fit to the data.

Finally, Chapter 7 discusses significant findings, theoretical and empirical contributions of the study, limitations, and opportunities for future research.

CHAPTER 2

LITERATURE REVIEW

In this chapter, I contextualize this research within previous scholarship. Figure 1 summarizes the structure of the chapter by indicating the main title of each section and subtopic. The chapter moves from a broad review of the primary studies on data, information, and knowledge sharing within organizations to scholarship on data sharing in public management. The first section offers a historical perspective of how researchers have conceptualized the role of information resources – knowledge, information, and data - in organizations over time. This evolution show why data sharing has gained a predominant role in current research, especially in public management. Then, I discuss the definition of "data" and I use the model developed by Spender in the late 1990s to describe data and information in public organizations. Spender's model is a useful lens to understand the unique characteristics of data and how these characteristics shape challenges and opportunities for data sharing within and across organizations. Finally, I provide an overview of the main theoretical approaches and empirical findings of current research on data sharing in the public sector. This final section sets the stage to introduce the theoretical framework developed in Chapter 3.

Figure 1. *Conceptual map of the literature review*

Information Resources Use in Organizations: Short Historical Perspective

Researchers have long been interested in how organizations and individuals acquire and process information resources – data, information, and knowledge (among others: Galbraith, 1973; Huysman & Wulf, 2006; March & Simon, 1958; Nahapiet & Ghoshal, 1998; Pfeffer & Salancik, 1978; Scott, 2003; Thompson, 1967). In early studies, organizations were considered closed systems of production, isolated from the external environment and primarily shaped by managers' decisions. Managers needed information only to maintain control over the organization and make decisions to improve efficiency and achieve organizational goals (Katz & Kahn, 1966; Pfeffer & Salancik, 1978; Selznick, 1949; Taylor, 1912). Information resources were internally produced, exchanged, and used, and researchers paid little attention to the exchange and sharing of information with external actors. Information, data, and knowledge were relevant as long as they provided insights on the organizational performance and internal outcomes, such as budget, finance, and production (Blau, 1955; Scott, 2003).

In the 1960s and 70s, researchers moved away from a closed system approach and adopted an "open systems" perspective. Open system researchers assume that organizations are highly dependent on their external environment, which "shapes, supports, and infiltrates organizations" (Scott, 2003, p. 29). Environmental factors include other organizations (Pfeffer & Salancik, 1978), sector characteristics (Hannan & Freeman, 1977), networks and communities (Inkpen & Tsang, 2005), and institutions and norms (Meyer & Rowan, 1977). Managers do not entirely control the organization performance because they cannot - or hardly can - influence environmental factors. Gaining access to information resources on external organizations, sector dynamics, norms and institutions, or networks became essential to maintain control over the environment (DiMaggio & Powell, 1983; Thompson, 1967). However, since other organizations control information, managers needed to interact and coordinate with them to access information. Open system researchers have increasingly focused on structures and incentives to facilitate resource access and exchange across organizations. These studies adopt a variety of theoretical perspectives, including social capital theory (Adler & Kwon, 2009; Chang & Chuang, 2011; Huysman & Wulf, 2006; Nahapiet & Ghoshal, 1998), organizational design (Galbraith, 1973), and network analysis (Chow & Chan, 2008; Inkpen & Tsang, 2005).

Public management research has also looked at how public organizations exchange data and information with stakeholders. Public organizations are information-intensive entities that collect and process large amounts of information to address the diversity of their stakeholder's needs, administrate a variety of services, and intervene in disaster and emergency management, among others (Bretschneider, 1990; Moon & Bretschneider, 2002; Willem & Buelens, 2007; Yang & Maxwell, 2011). Moreover, sharing of written internal documentation is fundamental to provide legal and technical justifications to actions performed by managers and administrators (Kornberger, Meyer, Brandtner, & Höllerer, 2017). In its well-known description of bureaucracy, Weber argues that the "public administration" is the ideal structure to manage and diffuse information across various organizational units (Weber, 1922).

There are several differences in how private and public organizations use and exchange data, information, and knowledge. Willem and Buelens (2007) find that the incentive structure of public organizations reduces the effectiveness and frequency of knowledge sharing across employees. Bozeman and Bretschneider (1986) show that information systems are different in public organizations as public managers need data to evaluate organization outputs not only in terms of economic efficiency but also equity, fairness, and effectiveness. Moreover, public organizations are characterized by tensions between a closed bureaucratic structure aimed to protect the technical competencies of bureaucrats (Kornberger et al., 2017) and quests for transparency and openness coming from citizens and external stakeholders (Meijer, 2015).

Over time, public organizations have evolved from closed bureaucracies to open organizations that frequently exchange data and information beyond organizational boundaries (Kim & Lee, 2006; Yang & Maxwell, 2011). Public organizations face a growing need to share information to coordinate with external stakeholders (Ansell & Gash, 2008; Hale, 2011; Kim & Lee, 2006; Koppenjan & Klijn, 2004; Saidel, 1991).

Moreover, pressures for transparency and accountability from civil society actors have pushed organizations to expand access to government information (Bertot et al., 2010; Kim & Lee, 2012; Romzek & Dubnick, 1987). According to Meijer (2015), the openness of public organizations is the result of political and societal development such as changing power relationships, new political actors and arenas, and technological and socio-cultural changes.

In response to these changes, managers and practitioners have increased their attention to knowledge, information, and data practices (Al-Alawi, Al-Marzooqi, & Mohammed, 2007; Kim & Lee, 2006; Willem & Buelens, 2007) and a growing amount of research has investigated the role of information in various policy domains, from emergency management to health and environmental policies (Comfort, 2007; Gil-Garcia & Sayogo, 2016; Kim & Lee, 2006; Klompas et al., 2012; Koliba, Zia, & Lee, 2011; Roberts, 2011; Vest & Issel, 2014; Welch, Feeney, & Park, 2016; Yang & Maxwell, 2011). An area of interest in the latest years is data use and exchange (Allard et al., 2018; Jennings & Hall, 2012). Data are the smallest unit of information available to managers and organizations, and provide novel information and knowledge to coordinate with other organizations, evaluate organization performance, and increase competitiveness. Researchers and practitioners are interested in how public organizations can acquire, analyze, and utilize data in their activities (Jennings & Hall, 2012). In the next section, I expand on the concept of 'data' and its unique characteristics.

Data as Information Resources

Data represent the smallest type of information resources, the unit upon which information and knowledge are built (Machlup, 1983). Data consist of raw information that is not yet or is hardly elaborated and presents itself in the form of numbers, words, figures, recorded voices, videos, and pictures (Drake, Steckler, & Koch, 2004; Hess & Ostrom, 2003). Government agencies use data such as financial figures, geographical coordinates, police recordings, tax returns, and census data. Public managers extract information from data in the form of observed patterns and trends. Information can be represented in visual forms, such as graphs or tables, or in texts, such as reports and presentations. The combination and integration of multiple information create knowledge regarding a specific topic (Davenport & Prusak, 1998; Tsai & Ghoshal, 1998).

Because of their raw format, data have a great potential to provide novel information to organizations. Organizations and individuals can elaborate, aggregate, and analyze data in different ways, thus extracting unique information and knowledge that can feed innovation and discovery and provide a competitive advantage (Nahapiet & Ghoshal, 1998). The importance of data as an input for new knowledge and innovation explains the growing interest in technology to analyze real-time, diverse, and granular data, often referred to as Big Data (Jennings & Hall, 2012; Kitchin, 2014; Townsend, 2013). Organizations and managers hope that the availability of data along with tools for managing and analyzing them will help produce new policy solutions and support decision-making processes.

The distinctive features of data compared to information and knowledge explain challenges and barriers to its exchange and use. In a widely cited model (Hess & Ostrom, 2003; Nahapiet & Ghoshal, 1998; Yang, Pardo, & Wu, 2014), Spender (1996) classifies information resources along two dimensions - tacit or explicit and individual or social. Tacit information resources are practices, policies, or skills that are not easily codified nor can they be explained to other individuals; therefore they are usually transmitted by example or long-term interactions. Instead, explicit information resources are codified into a structured language that can be understood by the individuals who receive the information. Individual information resources are controlled and recalled by a single individual, while social information resources are shared among individuals within a community and "reside in the tacit experience and enactment of the collectivity" (Adler & Kwon, 2002, p. 247).

Research has shown that tacit and social information resources are more difficult to share (Anssi, 2008; Hau, Kim, Lee, & Kim, 2013) because they require the ability to explain, integrate, and appropriate concepts that groups of individuals have mastered through experience and long-term interactions (Black et al., 2003; Bock, Zmud, Kim, & Lee, 2005; Willem & Buelens, 2007). Social and tacit information resources are deeply embedded in the relationships across individuals, so that organizations need to develop close relationships, common goals, and trust among their members if they want to facilitate the sharing of tacit and social information resources (Inkpen & Tsang, 2005; Leana & Van Buren, 1999; Nahapiet & Ghoshal, 1998; Tsai & Ghoshal, 1998). By contrast, individual and explicit information resources are more easily shared by individuals and organizations through standardized practices. For instance, organizations can hire more skilled managers or develop information systems that facilitate the transfer of codified information.

Table 1 illustrates Spender's model and provides examples of the different types of information resources in the public sector. Examples are drawn from Yang and Wu (2013) who illustrate the types of information resources more commonly used and shared by government agencies. Individual-explicit information resources include facts, data, and information that public managers and employees can consciously recall. Social-explicit information resources include value-added information and administration-oriented information. Value-added information is directly extracted by raw data that have been minimally elaborated or aggregated to fulfill the organization's information needs. It might include summary statistics or indicators that city departments have created to monitor their performance or illustrate policy problems. Administration-oriented information are government documents, meeting minutes, and notifications on work activities. They are "signals that connect government agencies to run their daily operations appropriately" (Yang & Wu, 2016, p. 33). Both value-added information and administration-oriented information are explicit because they are codified and easily transferable across organizations, and social because they reflect organization practices, values, and policies.

Administration-oriented and domain-oriented knowledge are tacit information resources. Administration-oriented knowledge includes best practices and abilities related to the management of public organizations. Domain-oriented knowledge refers to the experience gained and best practices learned regarding a specific policy field. Administration- and domain-oriented knowledge can be social if they refer to organization skills and practices, or individual if they refer to the professional experience that public managers have accumulated by working in the public sector or within a specific policy field.

Table 1.

Type of information resources in public administration.

	Individual	**Social**
Explicit	Facts, data (i.e. GIS data)	Added-value information (summary statistics, indicators) & administrative information *(minutes, reports)*
Tacit	Administration-oriented & domain-oriented knowledge *(professional experience and skills)*	Administration-oriented & domain-oriented knowledge *(organizations' practices)*

Adaptation from Spender (1996) and Yang and Wu (2013)

Research Boundaries

This research defines data as both raw and added-value information utilized by city departments to conduct their daily activities. The distinction between raw and value-added information is blurred, and it is rarely applied in practice and research (Weitzman et al., 2006). Public managers utilize the term "data" to refer to financial statistics or budget aggregated numbers, which are added-value information in Spender's model. Structured datasets that contain raw data should also be considered added-value information as they are designed based on the organization's information needs. Yet, they are usually called "data" by public managers and employees. Finally, the distinction between social and individual information resources is not relevant in this study. I assume that data are an explicit type of information resources that public managers can share through technical infrastructure and hardware supports (e.g., USB keys, CD-ROMs, hard disks) in contrast with knowledge that requires frequent, in-person, interactions to be shared.

More specifically, this research focuses on "administrative data" – i.e., data that public organizations generate and use in their daily activities. Administrative data include business licenses, vehicle registrations, tax records, birth and death records, or nonprofit activity records, among others. A recent report from the Pew Charitable Trusts (2018) shows that public managers in all US state governments engage in the collection, analysis, and use of administrative data to improve government action and effectiveness. Administrative data provide a critical input for the development of novel information and knowledge; are easily compared across departments and agencies; and can be easily shared with the public to enhance accountability and transparency (Allard et al., 2018; Blau, 1955; Jennings & Hall, 2012; Yang & Wu, 2016).

Public managers report that it is challenging to access administrative data beyond mandatory reporting and legal requirements (Pew Charitable Trusts, 2018; Allard et al., 2018). Because data are raw inputs, their value cannot be fully understood beforehand, and organizations might feel uncomfortable in sharing data with the government. Government agencies can use data to evaluate an organizations' performance or criticize decision-making processes. Moreover, technical and institutional barriers, such as lack of capacity and low interoperability across systems, differences in organizational culture and conflicts, and low trust, negatively affect the likelihood of sharing data with government agencies. Public management research provides several insights on barriers and incentives to share data. In the next section, I discuss the antecedents of data sharing in public organizations.

Data Sharing: Theoretical Perspectives

Data sharing is the transfer of data from one organization to another (Tulloch & Harvey, 2007). It can assume several forms (Roberts, 2011), from a "take it or leave it" approach where organizations provide public free access to data, to narrow agreements between two partners – i.e., a data user and a data provider (Susha, Janssen, & Verhulst, 2017). Open data initiatives fall on the first extreme. On open data portals, public organizations provide a set of aggregate and de-identified datasets that any organization and individual can access and download (Attard et al., 2018). On the second extreme, there are transfers of data between two organizations, including cooperative agreements that specify rules to exchange sensitive and personal data, restrictions on use, and limits to third-party transfer and analysis (Susha et al., 2017).

Between these extremes, data sharing assumes various forms from small collaborative projects, where groups of organizations occasionally share data, to structured networks where several actors regularly share data and information to coordinate everyday activities (Dawes, Cresswell, & Pardo, 2009; Dawes, Gharawi, & Burke, 2012; Weitzman, Silver, & Brazill, 2006). For instance, Dawes and colleagues (2009) coined the term "public sector knowledge networks" (PSKN) to distinguish information networks from networks finalized to the delivery and provisions of public services. PSKNs are networks of public organizations whose primary purpose is to share information resources (Dawes et al., 2009; Eglene, Dawes, & Schneider, 2007). PSKNs can include nongovernmental and international organizations, and they can focus on a single collective problem or an entire policy domain (Dawes et al., 2012).

However, there are substantial differences between open data and other data sharing initiatives. Open data portals respond to transparency, accountability, and participation demands from external stakeholders, particularly citizens (Grimmelikhuijsen & Feeney, 2016). Data sharing initiatives respond to managerial needs to collaborate with other organizations, provide public services more efficiently, or access data to support decision-making processes and policy design. Data sharing initiatives are generally more complicated because organizations need to develop common agreements, negotiate conditions for using and sharing data, address privacy concerns - especially in the case of sensitive data -, and design common infrastructure (Chen & Lee, 2018; Dawes et al., 2009; Gil-Garcia, Pardo, & Burke, 2010; Susha et al., 2017). Moreover, data sharing initiatives require significant financial and human resources to structure networks, coordinate data sharing, and integrate managerial and technical systems to transfer data.

In the past 20 years public management research has spanned from investigating the antecedents of successful data sharing partnerships and networks (Dawes et al., 2012; Gil-Garcia & Sayogo, 2016), willingness and frequency of data sharing within organizations and local governments departments (Bellamy, 6, Raab, Warren, & Heeney, 2008; Dawes et al., 2009; Welch et al., 2016) to employees' perceptions of confidentiality and trade-offs between transparency and privacy (6, Bellamy, Raab, Warren, & Heeney, 2007). The next section outlines the most common theoretical approaches and main findings of data sharing research. The literature review shows a progressive shift from a rational approach focused on costs and benefits to a socio-technical perspective which includes institutions and relationship characteristics. Integrative frameworks increasingly combine multiple approaches to explain data sharing within and across organizations.

Rational choice perspective on data sharing. A rational choice framework argues that organizations engage in data sharing based on expected costs and benefits; organizations and individuals share data if they expect that the total costs will be lower than the total benefits (Dawes, 1996; de Montalvo, 2003; Gil-Garcia et al., 2010; Nedović-Budić & Pinto, 1999; Tulloch & Harvey, 2007). Costs can include the implementation of technical infrastructures, adoption of new tools, loss of autonomy, time and human resources spent in negotiations, and managerial efforts (de Montalvo, 2003; Nedovic-Budic & Pinto, 1999). Benefits can include economies of scale, availability and diversity of information, and efficiency gains.

Research drawing from a rational choice framework tends to emphasize managerial and organizational factors that affect the perceptions of costs and benefits, including perceived barriers, organizational resources and capacity, and organizational culture and structure.

For instance, Dawes (1996) investigates data sharing barriers in the first data sharing initiatives across public agencies and she suggests three types of costs and benefits that managers evaluate before sharing data: (1) political – i.e., external influence, power, accountability and support; (2) organizational – i.e., self-interest, networks, and professional norms; and (3) technical – i.e., information infrastructures and data structures. As public managers report higher barriers across these categories, they perceive greater costs and are less likely to engage in data sharing with other organizations.

Other researchers have examined organizational resources and technical capacity arguing that organizations with lower financial and technical resources are less likely to engage in data sharing (Akbulut-Bailey, 2011). de Montalvo (2003) finds that public managers working in organizations with fewer resources do not expect to be able to engage in data sharing activities, perceive greater costs to increase their capacity, and therefore have lower incentives for sharing data. Similarly, Dawes (1996) and Gil-Garcia et al. (2007) argue that low capacity increases perceptions of costs while reducing the perceptions of benefits so as organizations are less motivated to share data.

Finally, some researchers focus on the organization culture, which consists of the values, formal and informal norms, and behavior patterns diffused within the organization (Schein, 1985). Public organizations promoting openness and collaboration are more likely to engage in data sharing than organizations promoting a proprietary, competitive culture and establishing strict control over data management (6 et al., 2007; Nedović-Budić & Pinto, 2000). In a competitive environment, self-interest prevails. Managers are more likely to perceive high costs from sharing data (Dawes, 1996) and prefer to protect their organization's position and assets rather than engage in inter-organizational exchanges (Gil-Garcia, Pardo, Baker, & others, 2007; Gil-Garcia & Sayogo, 2016; Pardo, Cresswell, Thompson, & Zhang, 2006). Vice versa, an organizational culture that promotes openness, public managers are more likely to perceive data sharing to be beneficial to organizational activities.

Overall, a rational choice approach emphasizes internal factors, such as organizational culture, structure, management, and financial resources. Researchers assume that public organizations and managers act as rational agents and share data with others when benefits are higher than costs. This perspective provided significant insights and led first researchers to investigate barriers and challenges affecting the implementation of data sharing initiatives. However, more recent work recognizes that rational approaches fail to consider non-instrumental sources of motivation (Fountain, 2007) and dedicate limited space to other social factors - institutions, relationships - which might shape data sharing.

Technology and social-technical perspective to data sharing. Technology has played a substantial role in the data sharing literature. As governments have adopted new technology tools, scholars have investigated different dimensions of technology, such as security and privacy, interoperability and standardization, and IT design (Sayogo, Pardo, & Bloniarz, 2014; Welch et al., 2016; Yang et al., 2014). Data is a structured and codified type of information resource, which it is easy to transfer, store, and retrieve with the support of technology tools such as hard disks, cloud storage, compact discs, and online platforms (Akbulut-Bailey, 2011; Bajaj & Ram, 2007). However, public managers need to manage concerns about security and privacy as well as differences in technical capacity across organizations. Public management research offers two perspectives on technology and data sharing; one draws from a rational choice approach while the second is based on socio-technical theories.

Early studies on technology drew from a rational choice approach (see section above) to argue that technical capacity is a necessary condition for organizations to lead successful data sharing initiatives. Technical capacity refers to the availability of IT infrastructures and tools along with an organization ability to implement, use, and manage them (Oliveira & Welch, 2013; Welch et al., 2016). Researchers find that high technical capacity greatly facilitates data sharing projects, but it is not a sufficient condition for data sharing (Gil-Garcia & Sayogo, 2016; Welch et al., 2016; Sayogo et al., 2014; Vest & Issel, 2014; Yang et al., 2014). Other social and institutional barriers become important as technical capacity is addressed within data sharing initiatives.

Interoperability across information systems is another key technical issue. Most organizations have information systems that are not or cannot be connected, hindering the opportunity to share data (Landsbergen & Wolken, 1998; Pardo et al., 2006; Vest & Issel, 2014). Interoperability can also refer to the standardization of data, including common formatting of shared data and metadata, the management of ontologies, and the harmonization of definitions and methods for collecting data across organizations (Gil-Garcia, 2010; Sayogo et al., 2014; Vest & Issel, 2014). There is heterogeneity in the formats that public organizations use to share data, “which include anything from images, PDF, and CVS files and Excel sheets, to higher structured XML files and database records” (Attard et al., 2018, p. 400). Lack of interoperability prevents public organizations from providing and receiving data; because data are organized differently, organizations cannot combine datasets and face challenges in extracting information from inadequate formats.

More recently, the study of technology in data sharing initiatives has moved beyond a rational approach to adopt a socio-technical perspective. Socio-technical theories argue that technology use and adoption are shaped by and depend on the social system in which technological tools, systems, and infrastructures are embedded (Bostrom & Heinen, 1977). The social system includes individuals, authority structure, organizational culture, and work characteristics, among others (Bostrom & Heinen, 1977; Feeney & Welch, 2014; Mehra, 2009; Oliveira & Welch, 2013). The impact of technology and its uses, applications, and outcomes are a function of how the social and the technological systems are integrated and how they interact with one another.

A socio-technical perspective emphasizes how the design of infrastructure and tools for sharing data is shaped by the characteristics of the social environment in which organizations are embedded (Bekkers, 2007, 2009; Black et al., 2003; Yang et al., 2014). Bekkers (2009) finds that public organizations organize information sharing in different ways. A "centralized integration" model occurs when one organization coordinates and re-distributes information to others. In an "interface connections" model each organization is responsible for sharing with others, while in an "information broker" model a third party facilitates and coordinates information sharing. Finally, a "shared information infrastructure" model is designed so as information is stored and uploaded in a shared space.

Different factors influence how organizations integrate information. For instance, asymmetries of power influence the characteristics of the information shared; authorization to access sensitive data; privacy and security solutions; and design of infrastructure, including the degree of integration and centralization of the system (Yang et al., 2014). Asymmetry in technological capacity can also affect infrastructure design (Vest & Issel, 2014). When public agencies have a relatively similar size and few functions, share power over common issues, and have a similar technical capacity, they are more likely to agree on a decentralized data sharing system in which no single organization is predominant (Yang et al., 2014). Centralized systems are more difficult to implement if organizations have similar power and size because no organization wants to delegate part of its autonomy to a central player.

Findings from socio-technical research challenge a rational choice approach to data sharing because costs and benefits are not the primary drivers of infrastructure design. Centralized systems are often less expensive and less complicated than decentralized ones, yet organizations prefer decentralized infrastructures (Yang et al. 2014). Overall, researchers note that technology is a necessary condition to increase government exchange, access and use of data, but technology is not sufficient by itself (Azad, 1998; Gil-Garcia et al., 2010). Organizations need to combine human, technical, institutional and social factors to understand the performance of data sharing projects and initiatives (Dawes et al., 2009, 2012; Roberts, 2011).

Inter-organizational relationships and data sharing. Studies on inter-organizational relationships include a variety of perspectives that investigate the characteristics of inter-organizational relationships including dependence and autonomy; trust and reciprocity; value and culture conflicts; and structures and incentives to overcome inter-organizational differences.

Resource dependency theory and inter-organization relationship studies suggest that data sharing depends on the characteristics of organization relationships, such as necessity, asymmetries, reciprocity, efficiency, stability, and legitimacy, as well as the mechanisms that regulate and support them (Azad & Wiggins, 1995; Oliver, 1990). For instance, organizations aim to maintain their autonomy. They engage in cooperative behaviors, including data sharing, only if the loss of autonomy is minimal or is compensated by greater outcomes and benefits for the organization (Azad & Wiggins, 1995). Moreover, organizations might be more likely to engage in data sharing if they face the necessity to coordinate with others, or the need to acquire legitimacy within their environment by engaging with other organizations.

Other studies highlight that past relationships across organizations affect data sharing, such as the willingness to provide data to common initiatives (Chen & Lee, 2018; Dawes et al., 2009; Gil-Garcia, Pardo, & Burke, 2010), and show that perceptions of trust, fairness, and equity in inter-organizational relationships are fundamental to achieve collaboration and data sharing (Black et al., 2003; Chen & Lee, 2018; Chow & Chan, 2008; Gil-Garcia et al., 2010; Nedović-Budić & Pinto, 2000; Karlsson, Frostenson, Prenkert, Kolkowska, & Helin, 2017).

Trust are positive expectations of others' behavior and a "willingness to be vulnerable" by relying on others (Rousseau, Sitkin, Burt, & Camerer, 1998, p. 394). Trust can develop through frequent and long-term interactions among organizations which shape expectations towards behaviors and attitudes of others (Black et al., 2003). For instance, trust increases expectations of reciprocity, which facilitate coordination, collaboration, and sharing of resources, because organizations expect that they will receive something in return and are more willing to invest resources into inter-organizational relationships. When actors lack or have low trust towards others, they might have negative expectations towards relationship outcomes and they need to implement strategies to prevent abuse and free riders, therefore increasing transaction costs, time, and effort required to support collaboration and resource sharing (Chow & Chan, 2008; Newell & Swan, 2000).

Karlsson and colleagues (2017) show that trust has a positive effect on inter-organizational information sharing in the public sector. Public managers can develop trust among parties involved by preventing the occurrence of events that undermine trust, such as information misuse or unmet expectations. Gil-Garcia et al. (2010) work echoes this finding and shows that cross-boundary information sharing relies on "trusted social networks," where past and present cooperation experiences among individuals are central to the develop of trust and, therefore, willingness to share data. Looking at a data network across public organizations in a metropolitan area in Nebraska, Chen and Lee (2017) find that public organizations who trust their partners are more likely to actively engage in the data network and collaborate to move forward the network goals. They note that trust develops either through frequent inter-personal interactions among managers, or through institutions, such as rules, data standards, and norms. Public managers who frequently interact develop a set of expectations regarding the behavior of other organizations, which increasing their willingness to rely on them, provide information, and collaborate. Institutions provide a set of shared expectations on how other organizations will behave because of their formal commitment to the network (trust stemming from institutions will be discussed to greater extent in Chapter 3).

Interpersonal relationships also help to overcome distrust and conflictual values among otherwise different groups, organizations, and departments (Chen & Lee, 2018). Relationships across organizations entail differences and similarities of values embedded in their culture (Drake et al., 2004; Sayogo, Pardo, & Bloniarz, 2014; Yang & Maxwell, 2011). As they collaborate, organizations need to deal with value conflicts that might hinder data exchange. Drake and colleagues (2004) discuss differences in values across three subcultures in the public sector - scientific, political and bureaucratic – which lead to different approaches to data and information use. Because of their training and professional beliefs, scientists mostly aim to analyze data and create reliable information. Politicians face the challenge to balance competing interests and they primarily utilize data and information to prioritize, deliberate, and justify decisions. Bureaucrats are concerned about legal requirements and consider data a commodity that feeds decision-making processes and organizational operations. Sub-cultures makes sharing data more difficult because of differences in values and expectations that lower trust and reciprocity. Furthermore, individuals differently organize data and information within datasets and indicators based on their sub-culture. Cultural differences might obstruct data use by other organizations which are not familiar with knowledge practices, values, and norms that are embedded into data (Atabakhsh, Larson, Petersen, Violette, & Chen, 2004).

Organizations can design structures and incentives to overcome difficulties stemming from the lack of trust or cultural differences. A project manager plays an important leadership role by coordinating the different actors and moderating cultural and organizational differences (Akbulut-Bailey, 2011; Black et al., 2003; Nedović-Budić & Pinto, 1999; Sayogo et al., 2014). Well-managed data sharing projects are more likely to successfully address challenges stemming from cultural clashes (Azad, 1998; Gil-Garcia & Sayogo, 2016). They are also more likely to secure funding and resources needed to successfully share information and data, which increase trust towards the project's ability to achieve its goals and provide incentives for organizations to participate (Chen & Lee, 2018; Fusi et al., 2018; Gil-Garcia & Sayogo, 2016). Finally, a project manager might facilitate the development of common data standards (Chen & Lee, 2018) which remove social and cognitive barriers that might prevent the exchange of information resources (Yang & Wu, 2016).

Other researchers emphasize the collaborative and voluntary nature of data sharing initiatives and focus on the conditions under which organizations provide their data. Conditions can include privacy terms and non-transfer agreements; responsibilities on shared datasets; limitations to commercial use or reproducibility; restrictions on scope and access; and so on. Susha et al. (2017) examine "data collaboratives", which are initiatives where private and public organizations cooperate to design incentives for the exchange and use of and access to a common pool of data. Data collaboratives face significant coordination problems and tackle them with a variety of instruments, ranging from negotiated agreements between two or more parties to unilateral agreements where the data provider establishes terms of data use and transfer without negotiating with the counterpart. Organizations can also adopt an intermediary model where a third-party organization facilitates the provider-user relationship, either by facilitating the negotiation of common conditions or matching providers and users that have similar needs (Attard, Orlandi, Scerri, & Auer, 2015; Susha et al., 2017; Zuiderwijk & Janssen, 2014).

In summary, literature investigating inter-organizational relationships in data sharing draws from a variety of theoretical backgrounds, ranging from social exchange and social capital theory to studies on network and collaboration. Findings consistently show that trust significantly facilitates data sharing and public organizations need to design managerial and collaborative structure that create incentives for sharing data or promote trust by regulating data use and transfer.

Institutions. The perspectives described so far - rational choice approach, socio-technical theories, and inter-organizational relationship studies - marginalize the institutional context in which public organizations are embedded and ignore "non-instrumental sources of motivation that stem from institutions" (Fountain, 2001, p. 65). Institutions are "humanly devised constraints that structure political, economic, and social interaction" (North, 1990, p. 3). Institutions shape organizational preferences and constrain organizational behavior and action by establishing common expectations, rules, and norms (DiMaggio & Powell, 1983; Fountain, 2001). The institutional environment has a strong influence on public organizations as they need to respond to legal and political pressures in order to maintain legitimacy among external stakeholders (Bozeman & Bretschneider, 1986; Frumkin & Galaskiewicz, 2004; Rainey, 2009).

There is mixed evidence regarding how institutions shape data exchange. Researchers suggest that formal institutions which mandate data sharing (e.g., rules, regulations, and organizational policies) facilitate data exchange within and across organizations (Gil-Garcia & Sayogo, 2016; Karlson et al. 2017; Landsbergen & Wolken, 1998; Yang & Maxwell, 2011). Formal institutions exert coercive pressures over public organizations and managers who are compelled to conform to these prescriptions. For instance, researchers find that Freedom of Information laws mandating government agencies to provide data upon requests from external stakeholders increase the provision of data and information compared to other written or informal requests (Worthy, John, & Vannoni, 2017). Public employees are more responsive and willing to cooperate with applicant requests when facing a legal obligation.

However, literature also suggests that bureaucrats might resist the mandatory provision of information. Bureaucrats can ignore institutional mandates due to resource constraints, internal divergences over the interpretation of the rules, or a lack of knowledge about rules and laws (Bauhr & Grimes, 2013; Worthy, John, & Vannoni, 2017). Institutional provisions might be in contrast with internal policies, other regulations protecting transparency and personal information, and, more broadly, with traditional bureaucratic values that emphasize internal expertise and processes over accountability and openness (Meijer, 2015). When institutions create uncertainty over the appropriate behavior, they leave space for public managers to act at their discretion.

The tension between openness and privacy is the focus of the work conducted by 6 and colleagues (6 et al., 2007; Bellamy et al., 2008) who investigate how public employees balance and align their actions with a web of formal and informal requirements; how formal requirements impact data sharing practices, particularly perceptions of confidentiality and trust; and the effect of both formal and informal institutions on interagency data sharing. Public employees face an imperative to share data and information with other organizations to improve public service provision and policy effectiveness. At the same time, they are required to comply with regulations and laws that address privacy and security concerns to prevent the release of sensitive information. 6 and colleagues find that formal institutions increase trust among agencies and managers' confidence in sharing data outside the organizational boundaries. Managers who work in highly institutionalized contexts are more confident that they understand other agencies' behavior and are able to judge when it is appropriate to share data. Managers also report that they rely on informal relationships and draw from their professional training and norms to deal with privacy concerns that are not addressed by regulation. Because regulation is still at its infancy, managers' initiative and discretion play an important role to bend or comply with formal rules.

Finally, Fountain draws attention to the institutional pressures that are internal to the public sector, such as the vertical structure of the bureaucracy, accountability, legislation, and budgeting. Public organizations are increasingly organized into horizontal forms of collaboration and interaction, such as networks. Yet, Fountain notes that the hierarchical and departmentalized structure of public agencies still hinder data sharing. In particular, formal and informal norms related to accountability, legislation, and budgeting procedures reinforce the separation of departmental activities and resources, institutionalize hierarchical responsibilities, and, overall, do not “fit” with the requirements of a collaborative and horizontal inter-organizational model of information and data sharing.

The study of institutions in data sharing is still at its infancy. Several findings are based on a small number of case studies which pay little attention to high-level institutions. Moreover, studies are generally limited to data sharing among public agencies, and we know little about how laws, regulations, and organizational policies affect cross-sector data sharing. Finally, hypotheses on institutional drivers of data sharing are often based on theory, but they have not been tested empirically.

Integrative frameworks for data sharing. As data sharing literature grows, researchers have combined theoretical approaches and empirical findings into integrative frameworks that provide a more comprehensive understanding of data sharing practices within and across government agencies. These frameworks depict the complex relationships among organizational, socio-technical, relational, and institutional factors and describe how they influence data sharing. Here, I highlight three main integrative theoretical frameworks that have been commonly applied in public management scholarship (6 et al., 2007; Fountain, 2007; Gil-Garcia, Pardo, & Burke, 2010; Yang & Maxwell, 2011).

Gil-Garcia et al. (2010) developed a framework entitled “Cross Boundary Inter-Agency Information Sharing” (CBIIS) to highlight the integration between technological components -such as technical infrastructures, interoperability, and shared data - and social ones - such as trusted social networks. The framework emphasizes the role of technology in data sharing while recognizing that technology implementation and adoption require trusted social networks among public managers. Trusted social networks buffer the conflicts that arise in the development of shared socio-technical systems and foster collaboration among organizations.

The strength of the framework is to highlight the critical technical components that government agencies need to consider while designing data sharing infrastructure. However, it falls short in including other social components, such as institutional and inter-organizational factors, that influence data sharing and infrastructure design (see previous sections on socio-technical perspectives and institutions) beyond the trust embedded in social networks.

Fountain (2006) integrated institutionalism with network theory and public management scholarship to propose a "multi-level framework for integrated information systems" (MIIS). The MIIS suggests that data flows are shaped by three types of factors: (i) formal institutions, including the legal framework, regulations, and contracts; (ii) organizational procedures, rules, and routines that shape public employees' behavior and transaction; and (iii) common practices and norms that are built by organizational groups and teams through collaboration. Data exchanges can occur at any of these levels and levels can mutually influence one another. For instance, high-level, formal institutions can shape organizational practices as well as preferences and incentives for the formation of groups and teams within organizations. Vice versa, relationships formed within groups and teams can modify institutions and design new organizational practices and norms for sharing data.

Fountain's framework tackles the complexity of data sharing and moves beyond the study of collaborative initiatives. Fountain considers data and information sharing as an essential component of the management of public organizations and an underlying theme to the adoption of new technologies. Her work explicitly theorizes the impact that the traditional structure and institutions of the public sector – e.g., bureaucracy and hierarchy - have on information flows. However, her framework does not address the challenges and barriers that public managers face when seeking data and information outside the boundaries of the public sector (Dawes et al., 2009, 2012). The framework also marginalizes the role of inter-organizational relationships and assign a limited role to technology infrastructure and design. Previous research has suggested that technology design depends not only upon internal culture and structure but also power and dependencies embedded in inter-organizational relationships (Yang & Wu, 2016).

Finally, Yang and Maxwell (2011) revised previous work on data and information sharing and proposed a set of models to explain data sharing within and across public organizations. They recognize that data sharing occurs at different levels - across individuals, across units, and across organizations - and argue that different factors shape data sharing within each level. At the individual level, public employees' attitudes toward data sharing and their perceptions of power and privacy shape data sharing. For instance, public employees with higher privacy concerns are less likely to share data. Across departments, data sharing is shaped by organizational culture and norms, social networks, reward structure, and power games, and common beliefs. Across organizations, technological, organizational, and political factors affect data sharing. Factors at different levels might be interrelated. For instance, social networks, trust, and individual attitudes might mediate the effect of legislatures and politics on data sharing. Within organizations, individual attitudes might shape the development of social networks, which in turn affect intra-organizational data sharing.

The value of Yang and Maxwell's literature review is to emphasize the different factors that researchers have investigated so far. Moreover, they highlight how current research lacks an empirical test of the indirect relationships that might occur across organizational, political, institutional, and technological factors.

Summary and Conclusion

This chapter reviews the current literature on data sharing, starting from a broad overview of research on information use in organizations to the definition of information resource types – data, information, and knowledge – and the theoretical approaches currently applied in public management research.

I define "data" as raw information that is not yet or is hardly elaborated and presents itself in the form of numbers, words, figures, recorded voice, videos, and pictures. Data are a crucial input to organizations because managers can use data to develop novel information and knowledge and improve organizational outcomes. Public organizations have a growing interest in data as technology has provided new, cheaper, and more efficient tools to collect, manage, analyze and store data. Yet, the availability of data is still highly dependent on the ability of managers to access data from other organizations.

Data sharing is the transfer of data from one organization to another. It can assume several forms, from one-to-one exchanges of data to collaborative projects and structured networks. Researchers have approached data sharing from different theoretical perspectives: rational choice frameworks focusing on costs and benefits as well as studies discussing institutions, inter-organizational relationships, and socio-technical factors. Integrated frameworks for data sharing advance research by combining multiple theoretical perspectives to capture the multi-level structure and the complex relationships that influence data sharing. Results from the application of integrative frameworks suggest that a single theoretical approach might not explain the complexity of sharing data and information.

From the review of the literature, I identify four main gaps that public management scholarships should address to inform public managers on how to improve access to and sharing of data.

First, few studies have considered how public managers exchange data with internal and external stakeholders in their daily activities. Most research is influenced by a collaborative governance approach which assumes that organizations are willing to share data and information collaboratively. For instance, case studies mostly focus on collaborative initiatives that promote data sharing within a policy domain or among actors that have common interests - e.g., face the same policy problem. Collaborative initiatives emphasize trust, governance, and common problems and goals but give limited attention to conflicts, power dynamics, and tensions that might arise in inter-organizational relationships. In reality, it is reasonable to assume that most data exchanges occur outside structured relationships. In their day-to-day activities, public managers request data from organizations that might or might not share them because they have no or little incentives to share data. Yet few studies look at data sharing that occurs across loosely connected organizations (Jennings & Hall, 2012; Welch et al., 2016), and most of them examine only intra-organizational sharing (6 et al., 2007; Kim & Lee, 2006; Willem & Buelens, 2007).

Second, we know little about public managers' role in data sharing and how they negotiate and manage the exchange of data and information across a variety of stakeholders. Current research largely focuses on the management of data sharing initiatives, including the deployment of technology tools and infrastructure, long-term funding, project scope, timing, communication, and demonstrable progress. Management assumes that organizations develop common organizational structures and practices (6, 2004), which constitute the premise for integrating data sharing processes. However, most public organizations are unlikely to adopt integrated solutions to share data because of their limited capacity and financial constraints. Public managers more likely opt to balance a mix of formal and informal coordination mechanisms to achieve collaboration with external stakeholders (6, 2004; Guo & Acar, 2005). A key question that public management research should address is under what conditions formal and informal coordination mechanisms are successful and how public managers should leverage formal or informal coordination to access data from a variety of stakeholders.

Third, most studies consider data sharing within and across public organizations, but they rarely investigate data sharing across sectors. Qualitative studies have found that barriers to exchange data are higher when public organizations engage with nongovernmental organizations (Azad & Wiggins, 1995; Cuganesan, Hart, & Steele, 2017; Dawes et al., 2009). Research on cross-sector collaboration also provides evidence that relationships with nongovernmental actors are more challenging than same-sector relationships, and public managers adopt different strategies across their relationship portfolio (Esteve, Boyne, Sierra, & Ysa, 2013). In order to increase data availability and ultimately improve decision-making processes, we need to consider the whole portfolio of relationships of public organizations and how antecedents of data sharing might differ across stakeholders.

Finally, most contributions to data sharing literature are theoretical or utilize data from few case studies (Dawes et al., 2009, 2012; Fountain, 2007; Gil-Garcia & Sayogo, 2016; Roberts, 2011). Few studies use large-scale quantitative methods to understand data sharing in the public sector (Welch et al., 2016). While case studies offer rich data to build theoretical hypotheses, quantitative analyses might provide generalizable results to guide policies at a national and federal level.

In the next chapter I propose an “Integrative Framework for Data Access” (IFDA) in the public sector. The IFDA aims to contribute to current literature by focusing its attention to the main gaps identified in this section.

CHAPTER 3

INTEGRATED FRAMEWORK FOR DATA ACCESS IN THE PUBLIC SECTOR

Despite the several studies discussed in Chapter 2, there remains a persistent interest in how managers can establish practices to access data from other organizations and improve data sharing within their agencies and outside organizational boundaries. This research assumes that public managers are information seekers who search and aim to obtain data from other organizations (Allard et al., 2018; Jennings & Hall, 2012; Levin, 1991; Susha et al., 2017). As public organizations are more interconnected within broad nets of relationships, public managers face a growing number of challenges and pressures to obtain information from a variety of sources to feed decision-making processes (Graber, 1992; Lee, 2013). The term "data access" indicates an organization's capacity to timely obtain data from other organizations. If access to information can improve decision-making processes, then understanding the determinants and likelihood of accessing data is fundamental to improve public outcomes and service quality. Public organizations with greater data access capacity will have greater availability of information and data to improve their performance. Moreover, public managers who can access data promptly can rely on up-to-date information to guide decision-making processes.

This chapter describes the theoretical framework of this research entitled "Integrated Framework for Data Access" (IFDA). The chapter presents each dimension and component of the theoretical framework. First, it illustrates the vertical dimension, which includes institutions, social environment, and coordination mechanisms. Second, it presents the horizontal dimension, which accounts for the portfolio of relationships from which public organizations request data. Finally, it describes the research context which includes cities with populations between 25,000 and 250,000 in the US.

Introduction to the Integrated Framework for Data Access

Data access is of theoretical interest because it provides the opportunity to investigate how public organizations and managers deal with a portfolio of relationships in a loosely structured environment. As highlighted before, a collaborative approach to data sharing fails to acknowledge how public agencies can improve data access as part of their daily activities and routines. While collaboration and cooperation across multiple actors encourage data use and positively affect trust, data sharing is most often challenged by proprietary and security concerns regarding data use and misuse, and power dynamics among data owners that wish to maintain control over information and data.

The Integrated Framework for Data Access (IFDA) emphasizes how data sharing efforts are often initiated by public organizations that need to access data to make decisions, control their environment, and provide adequate services to their constituencies (Graber, 1992; Lee, 2013). The IFDA draws from an open system approach which considers organizations embedded in a wide web of loosely structured relationships with other organizations (Pfeffer & Salancik, 2003; Scott, 2003). Obtaining resources from these relationships is necessary for an organization to survive, but it requires effort as other organizations often lack or have low incentives to collaborate. For instance, organizations might resist providing data to public agencies because it might damage their activities or erode their competitive advantage. Public managers play a crucial role because they can utilize strategies to bend inter-organizational relationships and therefore increase data access capacity within organizations.

The IFDA includes both a vertical and a horizontal dimension as illustrated in figure 2. The "vertical" dimension integrates institutional theory, resource dependence theory, and collaboration studies and highlights the multi-level structure of the factors that influence data access. Previous studies have widely suggested the appropriateness of a multilevel approach to investigate data sharing (Fountain, 2007; Yang & Maxwell, 2011). At the highest level, the institutional and social environment influence the behavior of the organizations. Institutional theory suggests that regulations and norms provide incentives and rewards for organizations that adopt prescribed behaviors and actions (DiMaggio & Powell, 1983; Meyer & Rowan, 1977). Resource dependency perspective argues that the exchange of resources - such as data - is influenced by the social structure of the environment in which public organizations are embedded, particularly power relationships. Data is a strategic resource and power affects the decision of organizations to disclose or withhold information to public agencies (Graber, 1992).

At the lowest level, coordination mechanisms show how public managers structure dependencies and relationships with other organizations that are not subordinated to them by hierarchy or market mechanisms (Malone & Crowston, 1994; Susha et al., 2017). Coordination mechanisms explain how collaboration and resource exchange occur between two or more organizations. For instance, city government managers can develop formalized arrangements, leverage personal networks, or communicate with other managers to facilitate data exchange among organizations (Tulloch & Harvey, 2007; Willem & Buelens, 2007). Coordination can also refer to data sharing agreements to regulate control, responsibilities, and data ownership in data sharing projects or organizational routines to exchange data (Nedović-Budić & Pinto, 1999; Susha et al., 2017). The data sharing literature hints at coordination mechanisms but lacks a solid theoretical and empirical framework to examine them.

The "horizontal" dimension emphasizes the portfolio of relationships from which public agencies request data. Public organizations can request data from other departments in their organization, other public agencies, or nongovernmental organizations. Previous literature suggests that there are differences between intra- and inter-organizational relationships as well as between same and cross-sector relationships when it comes to collaboration, interaction, and data sharing (Daley, 2009; Dawes, Cresswell, & Pardo, 2009; Eglene, Dawes, & Schneider, 2007). Researchers have advanced hypotheses on how data exchange is structured across this portfolio of relationships, but they have not systematically and theoretically examined the validity and contingencies of those prescriptions. Understanding differences across stakeholder types is important to increase data access capacity in public organizations. When departments obtain data and information from outside organizations but do not share them internally, public organizations might end up duplicating their efforts to collect data. Similarly, if internal data exchange is efficient, but departments have little access to external information, there might be little new information to feed decision-making processes.

This research expands current studies by investigating the social and institutional aspects that characterize each relationship type. In the IFDA, the horizontal dimension intersects with the coordination mechanisms to explore how coordination mechanisms apply across the portfolio of relationships of a public agency and how different barriers might affect data access in each relationship type.

Figure 2 shows the IFDA. The study's core hypothesis is that data access (the dependent variable) will be explained by systematic differences in the vertical and horizontal dimensions, such as differences across institutional settings, social environment, coordination mechanisms, and the portfolio of relationships. The research focuses on three types of data access: data access from other departments in the city, data access from other public agencies, including federal, state or local agencies, and data access from nongovernmental organizations, including both nonprofit and private for-profit organizations.

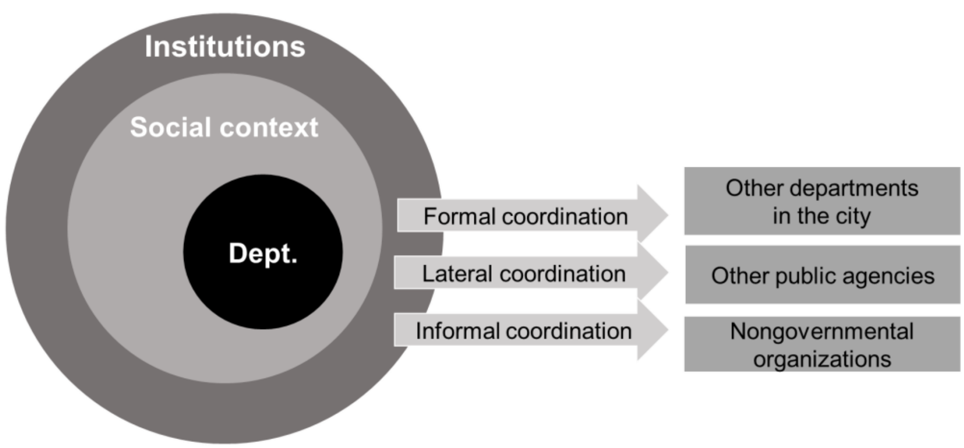

Figure 2. Integrative Framework for Data Sharing (IFDA)

Vertical Dimensions: Institutions, Organizational Context and Coordination Mechanisms.

The vertical dimension of the framework describes how and why institutions, social environment, and coordination mechanisms affect data access in the public sector. Each dimension is briefly described below.

Institutions. As discussed in the previous chapter, institutions are "humanly devised constraints that structure political, economic, and social interaction" (North, 1990, p. 97). The institutional environment in which organizations are embedded can include the political context, economics rules, culture, regulations, and laws, among others. Institutions can be either formal such as laws and regulations, or informal stemming from practices, norms, and values that regulate society, organizations, and groups.

Institutions transversally apply to all organizations and individuals, thereby shaping the framework within which organizations interact, exchange, and collaborate (Hannan & Freeman, 1977; Pfeffer & Salancik, 2003). Researchers who focus on institutions recognize that "organizations compete not just for resources and customers, but for political power and institutional legitimacy, for social as well as economic fitness" (DiMaggio & Powell, 1983, p. 150). Therefore, organizations show a tendency to conform to institutional pressures in order to be accepted by other actors, access resources, and collaborate.

There is a vast literature that investigates the role of institutions in collaboration and resource exchange. Studies on collaboration between public organizations and nongovernmental actors show that the institutional environment legitimizes collaboration and models the social, legal, and political structure that facilitates or hinders interactions across organizations (Allard et al., 2018; Bryson, Crosby, & Stone, 2006; Daley, 2009; Dawes et al., 2009; Smith, 2009).

Particularly, institutions provide a reference framework for organizations when they need to adapt to new structures or are forced to change internal procedures and routines (Dawes et al., 2009; Tolbert, 1985). Dawes and colleagues (2009, p. 398) note that “without an enabling policy framework, the risk-averse culture of government is likely to dominate decisions and actions.” Institutions enable the adoption of new practices and facilitate change within public agencies by defining non-instrumental incentives (Fountain, 2001). In some cases, institutions explicitly codify changes and mandate inter-organizational collaboration via formal norms and rules. In some others, institutions are used to signal the behaviors that are socially accepted in the environment by other organizations.

Institutional theory has also suggested that institutions might play an important role in trust-building processes. Research has widely shown that trust among actors is one of the main antecedents of data sharing and collaboration (Chen & Lee, 2018; Dawes, Cresswell, & Pardo, 2009; Gil-Garcia & Sayogo, 2016; Thomson & Perry, 2006). As mentioned above, trust is the “willingness to be vulnerable” and to take risks by interacting, exchanging resources, and collaborative with other actors. Trust stems from frequent interactions, whereas actors interacting more often are more likely to develop positive expectations about others. Yet trust developed within a dyadic relationship benefits only the actors involved. Moreover, differences across actors and emerging conflicts might challenge trust in the long run (see: section “inter-organizational relationships and data sharing”).

Institutional studies note that when trust stems from institutions, it benefits not only the actors involved in the relationship but the broader context in which an organization operates (Smith, 2009). Social norms, for instance, encourage actors to maintain a certain behavior and make them more predictable, thereby increasing organizations’ willingness to “trust” him. A shared legal framework can define expectations and set limits to organizational action; under these conditions, organizations might be more willing to cooperate because institutions ensure the safety and legitimacy of the collaboration (Smith, 2009).

In the IFDA, I distinguish between coercive and normative pressures that institutions might exert over organizations. Privacy laws and legal mandates to share data are primary sources of coercive pressures over organizations because they codify expectations, create constraints that organizations need to comply with, and define national and local incentive structures to share data (DiMaggio & Powell, 1983; Berry & Berry, 2007). Formal institutions might either enable data sharing by establishing common rules for organizations and increasing trust in the environment, or hinder data sharing as organizations become afraid of violating legal prescriptions.

Normative pressures occur when institutions shape preferences and cultural expectations of organizations (DiMaggio & Powell, 1983; Fountain, 2007). For instance, open data research shows that citizens and external stakeholders might exert pressures for public organizations to provide their data online in an open access format (Grimmelikhuijsen & Feeney, 2016). Normative pressures are particularly effective when behaviors and actions of actors are highly visible to others. For instance, studies have shown that the adoption of a website and other e-government tools by some organizations has created normative pressures towards other organizations to adopt similar tools and conform to the expectations of the environment in order to maintain legitimacy (DiMaggio & Powell, 1983; Frumkin & Galaskiewicz, 2004; West, Mayer-Schönberger, & Lazer, 2007). In the case of data sharing, visible informal institutions might include websites promoting transparency and openness, task forces, strategic documents, or organizational position dedicate to data management (Pew Charitable Trusts, 2018; Tolbert, 1985).

Social environment. City departments are increasingly connected with one another and with other nongovernmental organizations in order to address the complexity of policy design and public service delivery (Hale, 2011; Koppenjan & Klijn, 2004). Social relationships are instrumental to organizations because they facilitate or hinder access to resources needed to achieve organizational goals (Levine & White, 1961). Research on data sharing suggests that relationships may vary by type, characteristics, and frequency of interactions in ways that promote or hinder the formation of data exchange relationships (Dawes et al., 2009; Gil-Garcia & Sayogo, 2016; for a complete review see chapter 2).

Power dynamics are an important predictor of data sharing. Power is defined as the capacity of one organization to pressure another one and orient its decisions to address the needs of its membership (Emerson, 1962, 1976; Molm, 1994). Resource dependency theory suggests that power is key to understand the nature of relationships and the flow of resources among organizations; organizations engage in transactions and exchanges with other organizations to either maintain, improve, or reinforce their power position (Pfeffer & Salancik, 1978). When external organizations hold power over public organizations, they might be more likely to refuse to share data and withhold the requested information. By contrast, public agencies in a power position might have easier access to data from other stakeholders.

In a recent article, Meijer (2018) highlights the data ecosystems in which cities are embedded and the power dynamics among involved actors. He describes examples of private companies, such as Airbnb and Uber, which refuse to provide data to public agencies. Power dynamics reflect the value associated to data access and use as well as political goals of actors involved (Garcia, Pardo, and Nam 2016). Data are raw inputs that might provide organizations with novel information and knowledge to improve their performance and gain a competitive advantage over other organizations. They can also be manipulated and used to support positions or strategies that damage the data owner (Meijer, 2018; Meijer & Thaens, 2018; Stone, 2001). As they represent such critical resources, data exchanges are characterized by issues of power across organizations.

Coordination mechanisms. Current scholarship lacks a systemic investigation of coordination mechanisms for data sharing. Collaboration and network management studies suggest that government managers need to balance the integration of formal institutional structures, such as memoranda of understanding, with informal coordination based on personal relationships and commitment (Chen & Thurmaier, 2009; Thomson & Perry, 2006). Based on these studies, the IFDA recognizes that managers design and implement diverse coordination mechanisms to exchange data with other stakeholders and test their effectiveness across the portfolio of relationships.

Researchers have classified coordination mechanisms according to different dimensions. March and Simon (1958) differentiate between coordination "by plan" when common activities are managed in advance and coordination "by feedback" when common activities are managed ex-post. Malone and Crowston (1994) note that coordination might address different types of interdependence across organizational tasks and operations: prerequisite interdependence, shared resource, and simultaneity. Mandell and Steelman (2003) differentiate between intermittent coordination through temporary task forces coordinating specific tasks or domains, and permanent coordination through formal arrangements.

In data sharing literature, several researchers have used formalization as the discriminant factor to distinguish among coordination mechanisms for exchanging data, information, and knowledge (Guo & Acar, 2005; Weitzman, Silver, & Brazill, 2006; Willem & Buelens, 2007). Formalization refers to the extent to which coordination is implemented and adopted at the organizational level through written agreements. At one extreme, organizations develop formalized routines to exchange information resources; on the other, public managers exchange data through personal relationships. Roberts (2011) finds that data exchange in the public sector occurs through either personal relationships among public managers and employees or structured coordination if organizations develop rules to control the data sharing process, plan activities, and establish formal roles. Similarly, Weitzman and colleagues (2006) note that data exchange mechanisms "can be positioned along a continuum of formality from highly structured, formalized agreements to systematically collect information to informal data-sharing relationships among agencies or analysts within agencies" (p. 391).

The IFDA draws from the work conducted by Willem and Buelens (2007) who classify coordination mechanisms in formal, lateral, and informal coordination. This classification includes most mechanisms that have been discussed in previous research (Dawes, 1996; Gil-Garcia & Sayogo, 2016; Kim & Lee, 2006; Tulloch & Harvey, 2007; Welch et al., 2016; Yang & Maxwell, 2011). Formal coordination is based on written agreements among organizations to design routines to regularly exchange data and information. Routines can take the form of regular exchanges of data via email; submissions of data through IT systems; or scheduled uploads of data on a cloud platform. Lateral coordination relies on written agreements that are developed between two or more organizations on a case-by-case basis. For instance, public managers can sign a data sharing agreement to provide a dataset to another organization or submit a Freedom of Information Act (FOIA) request to obtain data from another public agency. Compared to formal coordination, lateral coordination does not assume that the exchange of data will repeat in the future and does not lead to regular exchanges of data among the organizations. Finally, public managers can access data through informal coordination by leveraging personal and professional relationships. They can contact previous coworkers, friends, and other personal and professional connections in other organizations in order to obtain data.

Horizontal Dimension: Relationship Portfolio

Government agencies have become increasingly dependent from external stakeholders due to the growing complexity of the policy environment, globalization, reform movements to limit government, and advancements in technology (Fishenden & Thompson, 2013; Guo & Acar, 2005; Koppenjan & Klijn, 2004). For instance, government agencies are partnering with nongovernmental organizations to increase efficiency and effectiveness of public services and develop new solutions in the face of significant budget cuts, particularly at the state and local level (Malatesta & Smith, 2014). Other government agencies are privatizing or outsourcing public services to save financial and human resources (Alonso, Clifton, & Díaz-Fuentes, 2015; Avery, 2000; Girth, Hefetz, Johnston, & Warner, 2012; Marvel & Marvel, 2008) or working with stakeholders to address complex policy problems (Koppenjan & Klijn, 2004; Lubell, Schneider, Scholz, & Mete, 2002) and intervene in disasters and crises (Kapucu & Hu, 2016; Wukich, Siciliano, Enia, & Boylan, 2017).

As government agencies more extensively work with external stakeholders, they increasingly need to share data with a variety of external and internal stakeholders to coordinate common activities, oversee goals, assess the performance, and target the provision of public services, among others (Jennings & Hall, 2012). The diversity of a public agency's stakeholders impacts data access because of the different institutional and social factors that characterized intra- vs. inter-organizational relationships and same- vs.

cross-sector relationships (Daley, 2009; Yang & Maxwell, 2011). While extensive and varied networks might provide access to more valuable data, their complexity can be difficult to manage because of cultural and cognitive differences, goal conflict, technical barriers, and power issues (Dawes, 2009).

Previous research shows that departments in the same city share norms, culture, and values which facilitate the exchange of information resources (Bolino, Turnley, & Bloodgood, 2002; Tsai & Ghoshal, 1998). Similarities in professional norms also facilitate networks of public organizations, but the bureaucratization of managerial activities, jurisdictional boundaries, and task specialization obstacle data and information sharing (Daley, 2009). Differences in values and professional norms are particularly stark in cross-sector relationships, where public managers need to remove cognitive and cultural barriers that might hinder data access (Ansell & Gash, 2008; Bryson et al., 2006; Daley, 2009; Guo & Acar, 2005).

This research explores differences across the portfolio of data sharing relationships by examining three types of stakeholders: other departments in the same city, other public agencies, and non-governmental organizations, including for-profit and non-profit organizations. Examining data sharing antecedents across stakeholder types is important for at least two reasons. First, we need to understand how data sharing is regulated within and across organizations to increase data access capacity within the public sector. If department heads effectively obtain data and information from other organizations, but they do not share with other departments in the city, public organizations might duplicate effort to access data or data access might provide benefits to only a few departments (Bate & Robert, 2002; Szulanski, 2000). Likewise, if public managers only share internally, they might face a shortage of information about the environment or have access only to redundant information across departments.

Second, research has little explored factors that are unique to each context and factors that are common across the relationship portfolio (Yang & Maxwell, 2011). From a theoretical perspective, combining research on intra- and inter-organizational data sharing can provide new insights into how data sharing is structured. Empirically, results can help public managers to identify effective ways to increase data access.

Application of the Integrative Framework for Data Sharing: Data Access in Small and Medium Sized Governments.

I apply the IFDA to investigate data access across US small- and medium-sized cities. Data access is an essential part of public managers' work but obtaining data from other organizations might require significant time, efforts, and resources (Allard et al. 2017). Large cities are likely to have the resources to lead the adoption and implementation of technologies to collect and share data and to bend formal rules that might hinder data sharing by creating new partnerships and collaborative initiatives (Jimenez, Mossberger, & Wu, 2012). Instead, small- and medium-sized cities often lack both the organizational and managerial capacity to lead inter-organizational consortiums and have to rely on individual entrepreneurship to mobilize resources (Fusi & Feeney, 2016). Moreover, previous studies show that private interests and local power groups significantly influence policymaking in small cities, increasing the likelihood that public managers will face power dynamics and tensions in accessing data and information

(Hamin, Gurran, & Emlinger, 2014; Hamin et al., 2014). Finally, small- and medium-sized cities benefit less from digital activist groups that promote citizen-driven open data and crowdsourcing initiatives that increase data availability in large cities (Graham, Hogan, Straumann, & Medhat, 2014). These differences across small and large cities make it particularly important to investigate data access in this research setting.

CHAPTER 4

HYPOTHESES

Chapter 4 outlines the hypotheses of the study. According to the IFDA, data access is a function of institutions, social environment, and coordination mechanisms (vertical dimension) as well as stakeholder differences across the portfolio of relationships of a city department (horizontal dimension). This chapter applies the IFDA and develops hypotheses on the antecedents of data access in small- and medium-sized cities in the US. Figure 3 summarizes the empirical framework.

The chapter is structured as follows: first, I focus on coercive pressures stemming from privacy laws and legal mandate, as well as normative pressures stemming from state institutional capacity and visible signals, such as open data portals. Then I move to the social environment and discuss how stakeholders' influence impacts on data access. Finally, I discuss coordination mechanisms – formal, lateral, informal and technology tools – and their effectiveness across the relationship portfolio.

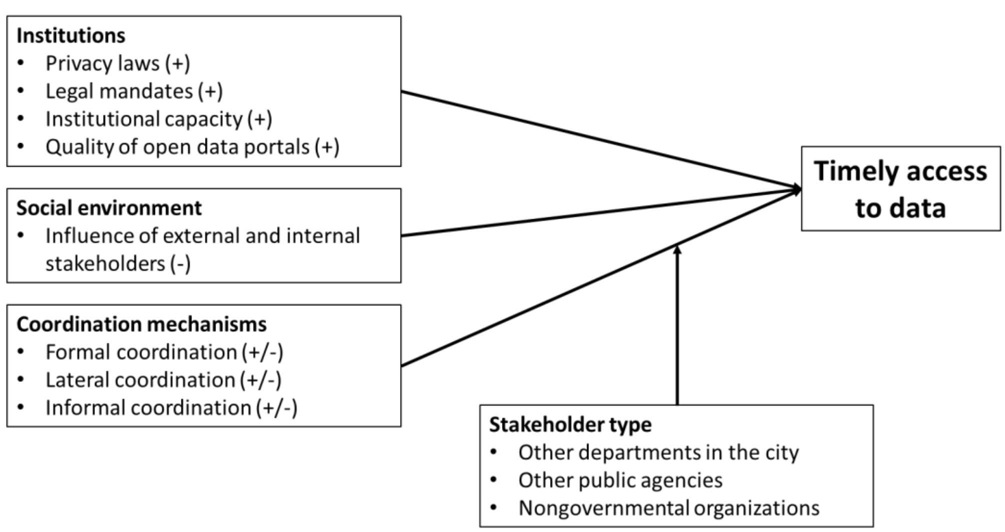

Figure 3. Summary of the hypotheses

Institutions

State governments are an important source of coercive and normative pressures for small- and medium-sized city governments, as they establish most of the policies, rules, and regulations that city governments need to comply with. State governments hold a position of power over city governments and can set expectations, provide symbolic incentives, and influence organizational decisions. For instance, state governments provide financial resources – grants, subsidies – which create incentives to comply with normative pressures or legally force organizations to comply with funding rules. Here I consider both coercive pressures stemming from state-level privacy laws and legal mandates and normative pressures stemming from state-level institutional capacity and signals of transparency.

Coercive pressures: Privacy laws. Privacy laws bind organization actions and behavior by establishing rules to share and use data. Through privacy laws, state governments aim to balance the need for sharing data across organizations to enhance organization performance, accountability, and transparency with the need for protecting the individual rights of their citizens.

In the United States, privacy laws are relatively weak as compared to other countries (Baumer, Earp, & Poindexter, 2004; Leuan, 2017; Steinhardt, 2010). There is no comprehensive federal law regulating privacy rights; the few existing legal provisions are contained in policy-specific bills, such the Health Insurance Portability and Protection Act and the Family Educational Rights and Privacy Act. Most privacy laws are proposed and enacted by state governments to regulate issues concerning data collection, use, and diffusion; third-party sharing of data; data breaches notification; individual right to privacy; and organizations' duties (Leuan, 2017). Because state governments establish privacy laws, the institutional landscape is highly fragmented. A study conducted by a US law firm finds that most US states have enacted only fundamental data-related laws to protect citizens from unauthorized access to their data, especially when data contain sensitive information. Few states explicitly recognize individuals' rights to privacy and protect the use and transfer of data ("Data Protection and Security Solutions for State and Local Government," n.d.).

There is mixed evidence on the effect of privacy laws on data access (Yang & Maxwell, 2011). On the one hand, privacy laws might increase data access. As noted in the previous chapter, a legal framework can act as a substitute for trust as it establishes minimum requirements and expectations that parties involved are required to respect (Smith, 2009). For instance, organizations providing data to city departments might be afraid of data breaches, unauthorized use of data, or other security issues. External stakeholders often perceive governments as lacking capacity to manage the security of data and information; stricter privacy laws might reduce such concerns as they establish minimum levels of security and plan actions that governments are required to implement. When state governments provide clear indications about privacy rights and limits, they influence trust across organizations and facilitate the exchange of resources (Fountain, 2001; Smith, 2009).

Moreover, privacy laws might increase perceptions of legitimacy as they authorize the sharing of data and information. The scope of privacy laws is to establish which data and information organizations can or cannot share. Previous studies find that employees often refrain from providing data to other organizations because they perceive data sharing as risky behavior and are worried about negative consequences that can occur by disclosing organizational information (6, 2004). Explicit legal provisions increase confidence and awareness across employees on whether they can provide data and information to third parties, thereby increasing the likelihood that they will share data.

On the other hand, privacy laws might prescribe specific limitations hindering data sharing. In some cases, privacy laws prevent organizations from transferring data to third parties or limit the type of information that organizations can share. In a recent study conducted by the Pew Charitable Trusts, a manager from a state government agency complains that privacy laws prevent the integration of databases across public organizations because they ban the use and sharing of individual identifiers (Pew Charitable Trusts, 2018). Privacy laws might also increase concerns about compliance with legal requirements. 6 and colleagues (2004) find that coercive institutions can increases costs and concerns among public employees who need to comply with process-based requirements. By establishing boundaries for sharing data and treatment of information, privacy laws place constraints and pressures upon organizations and managers who might be worried about conforming to privacy laws and refraining from exchanging data with other organizations. Privacy laws can be a major obstacle to data sharing projects, even when organizations are willing to share their data (Pew Charitable Trusts, 2018).

In the case of city governments, I argue that privacy laws are likely to provide incentives for external organizations to provide data to city departments. Local governments are at the center of discussions regarding privacy and security concerns (Newcombe, 2017; Norris, Joshi, Laura, & Tim, 2018). Given the low technical capacity of small cities, external stakeholders might be unwilling to share data without an adequate legal framework to protect data use and disclosure. Therefore, in states where privacy laws are stricter, external stakeholders might be more willing to collaborate with the local government as they feel more comfortable about data and information treatment. Privacy laws might reinforce the position of the local government and facilitate data access. Therefore, I expect that:

H1a: State privacy laws will be positively correlated with data access.

Coercive pressures: legal mandate of sharing data. While privacy laws provide a general framework for sharing data, state governments can force organizations to provide data to government agencies and departments. State governments often mandate data sharing to facilitate the implementation and design of public policies; evaluate organization activities and obtain benefits and funding from the government; or comply with regulations concerning reporting and accountability. Nonprofit organizations, for instance, are required to provide their financial information to government agencies, such as the IRS, to demonstrate compliance with the nonprofit regulation. City departments are required to share information with police departments for matters related to security and order.

Legal mandates exert strong coercive pressures; organizations are willing to comply with legal prescriptions to avoid fines and penalties. Several researchers find that legal authority is an important antecedent of data and information sharing (Chen & Lee, 2017; Dawes, Cresswell & Pardo, 2009). Most government agencies report that data sharing mandates play a fundamental role in ensuring data access from external stakeholders, both other public organizations and nongovernmental ones (Allard et al., 2017). In some cases, managers even acknowledge that state legal mandates are the most effective mechanism to access data from other organizations. Legal mandates simplify data access by establishing clear responsibilities and roles, ownership over data, and procedures for sharing (Pew Charitable Trusts, 2018).

Based on previous evidence, I expect that state legal mandates will be positively correlated with data access. Empirical and theoretical studies on data sharing and open government find that the adoption of laws that are supportive of data sharing has a positive effect on inter-organizational exchange of data and information (Pew Charitable Trusts, 2018; Piotrowski, Rosenbloom, Kang, & Ingrams, 2018; Worthy, John, & Vannoni, 2017). In small- and medium-sized local governments, legal mandates might be particularly crucial as public managers lack the institutional capacity to influence other organizations' behavior. Therefore, I hypothesize that:

H1b: A legal mandate to share data across organizations will be positively correlated with data access.

Normative pressures: institutional capacity. Normative pressures push organizations to conform to cultural norms. Organizations comply with normative pressures and adopt behaviors that align with the environmental context because they wish to maintain stable relationships with other organizations and gain access to resources (Jun & Weare, 2010; Frumkin and Galaskiewicz 2004; Meyer & Rowan 1977; Scott 2011). Normative pressures can push organizations to adapt even when institutions are in contrast with their structure and goals or do not lead to improvements in the organization performance (DiMaggio & Powell, 1983; Meyer & Rowan 1977).

Normative pressures are higher when cultural norms stem from actors that are in a position of power because they hold financial, material, and symbolic resources (DiMaggio & Powell, 1983) and, therefore, have greater influence over other organizations (Chen & Lee 2017; Gazley, 2008). State governments, for instance, provide significant financial resources to local governments and award legitimacy to local government activities. Cultural norms that are pushed forwards by state governments are likely very effective in influencing city governments' behavior.

State governments can dictate cultural norms by designing institutional arrangements that support sharing and openness of government data, and therefore building "institutional capacity" that guarantees government ability to accomplish its goals and support policy action (Peters & Pierre, 1998). By creating institutional capacity, state governments show commitment towards specific policy goals, promote the convergence of goals across actors, and set general expectations for city governments (Smith, 2009). For instance, research on e-government shows that state-level institutional capacity in the form of dedicated legislative committees or independent information technology executive departments positively affect the likelihood that cities will adopt new information technologies (Grimmelikhuijsen & Feeney, 2016; McNeal et al., 2003a). Studies on inter-organizational collaboration also show that when state governments commit human and financial resources to create infrastructures that facilitate coordination and interaction across city governments, public managers are more likely to collaborate (Smith, 2009).

Dawes and colleagues (2009) find that public managers perceive the lack of support from high-level government agencies to be a severe barrier to access data and information from other stakeholders. State governments can build institutional capacity for data sharing in several ways: they can allocate funding and resources to data sharing initiatives; establish a Chief Data Officer to manage data and data sharing policies; or design standardized data sharing agreements to help public managers comply with privacy laws and citizen rights. When state governments adopt practices, structures, and incentives that are conducive of data sharing, city governments might be more likely to report high data access as organizations will conform to the institutional environment (Chen & Lee, 2018; Yang & Maxwell, 2011). In fact, when state governments build institutional capacity to support data sharing, it is more likely that organizations will feel greater normative pressures to provide data to city departments. Therefore:

H1c: State governments' institutional capacity will be positively correlated with data access.

Normative pressures: Open data portals. Visible actions and social cues are another way through which state governments can promote and diffuse cultural norms.

Organizations rely on social cues and inputs from the environment to understand which behaviors and actions are acceptable by others.

Signals from the environment can guide public organizations and managers when they have to interpret data use and sharing rules. 6 and colleagues (2014) note data sharing rules are often ambiguous and do not apply to all concrete cases that public managers encounter. Therefore, public managers rely on their interpretation of the ethics and rationales underlying the regulation to fill legislative gaps. Institutional theory suggests that public managers' interpretation is likely influenced by cultural norms that guide expectations and acceptable behaviors among actors in the same environment. If cultural norms are favorable to openness and transparency, public managers are more likely to align with them by providing data to other organizations.

I suggest that state governments can set cultural norms and shape expectations on data sharing by promoting open data portals. Open data portals are highly visible social cues signaling a state government's commitment towards transparency and accountability (Jun & Weare, 2011). Open data portals vary significantly across state governments. Some states have put considerable efforts into building and promoting open data, setting an example for local governments and other organizations. Others provide little information online, limit access to data, or offer only incomplete, not up-to-date data to the public.

Because open data portals are social cues that indicate commitment towards transparency and accountability and signal acceptable norms and behaviors across organizations, I suggest that in states where open data portals are of higher quality, city departments will enjoy greater data access from other organizations. Quality – such as the extensiveness and accessibility of open data – signals commitment from the government to a culture of transparency which will positively affect relational trust across organizations and will promote an interpretation of existing rules and regulations that favors data sharing. Therefore:

H1d: Quality of state government open data portals will be positively correlated with data access.

Social Environment

The social environment incorporates the dynamics of power that characterize the relationship between a city department and its external stakeholders. Power is defined as the ability of an organization to pressure another organization to address its own interests. Power dynamic are widely addressed by resource dependency theory (Pfeffer & Salancik, 1978) and recent studies in public management that recognize data access as an issue of power among organizations involved (Meijer, 2018).

Influence. City departments are part of loosely structured policy systems in which they share power over decision-making processes with a multitude of stakeholders, such as local committees, representatives of the civil society, and interest groups (Moore, 1995). Stakeholders receive representation power and mandate from their membership and constrain a city department's activities by exerting pressures towards the goals and interests they represent (Dawes, 1996; Frumkin & Galaskiewicz, 2004). The degree to which stakeholders can influence a city department's activities varies according to the material and symbolic resources that they control, including funding, reputation, advocacy, legitimacy, and political support (Saidel, 1991; Suárez, 2011).

City departments have an interest in accessing data from other organizations, especially when these organizations hold great influence over decision-making processes. First, influential stakeholders are more likely to attract the attention of city departments because they participate in and shape policymaking processes (Pfeffer & Salancik, 1978). Second, city departments are interested in collecting data and information from influential stakeholders; data and information might help public managers to balance information asymmetries, gain more control over the policy environment, and increase their centrality in the policy discussion. However, collaborating and exchanging resources with influential stakeholders is more difficult, and public managers face a greater likelihood of conflicts (Simo & Bies, 2007).

Previous research has shown that power dynamics negatively affect data exchange across organizations generally and within the public sector specifically (Dawes, 1996). Pardo and colleagues (2006) find that regulatory agencies refrain from sharing data because they fear losing relevance and power in their policy domain. Similarly, Gil-Garcia and colleagues (2005) and Azad and Wiggins (1995) observe that public agencies consider the relative influence of other organizations over their activities before engaging in data sharing with them. If public agencies believe that external organizations already have too much influence on policymaking, they will be less likely to share data. Once organizations share data with others, they can hardly control how data are used (Blau, 1955). Organizations receiving data might utilize them in ways that damage the original data owner, for instance by creating new information and knowledge that allow them to gain a competitive advantage or increase control over the environment. Providing data to the city departments might change an organization's relative position and increase the city influence over them (Yang & Maxwell, 2011). The collaboration literature confirms similar dynamics across organizations in different sectors, where influence and power negatively affect the exchange of resources and collaboration (Bryson, Crosby, & Stone, 2006; Simo & Bies, 2007; Thomson & Perry, 2006).

Given the evidence from previous studies, I expect that increases in the level of influence of external organizations will negatively affect data access, while city departments that experience lower levels of external influence – i.e., they are central in their policy environment – will report greater data access.

H2a: Influence of external organizations will be negatively correlated with data access.

Coordination Mechanisms & Relationship Portfolio

Public managers apply a mix of coordination mechanisms to increase the likelihood to access data. The horizontal dimension of the IFDA suggests that the effectiveness of coordination mechanisms vary across the portfolio of relationships. This insight draws from a large body of literature suggesting that there are sector-based institutional and social differences that public managers need to consider when exchanging data (Daley, 2009; Roberts, 2011). The next paragraphs explore three forms of coordination mechanisms – formal, lateral, and informal – and how they apply to data access from other departments in the city, other public agencies, or non-governmental organizations.

Formal coordination. Formal coordination – e.g., the development of written data sharing agreements and routines – has received much attention in data sharing literature. Researchers find that formal coordination facilitates data exchanges by explicitly addressing political and security risks as well as social and technical barriers that commonly hinder data sharing (Dawes et al., 2009). By formally coordinating data exchange, organizations agree on common rules and data standards, design shared policies, establish responsibilities and rules, and distribute benefits and costs among parties involved in the exchange (Dawes, Cresswell, & Pardo 2009; Landsbergen & Wolken 2001; Simo & Bies 2007).

Formal coordination, however, entails lengthy negotiations among partners to agree on sensitive issues such as confidentiality, privacy, responsibilities, and costs (Allard et al., 2018; Dawes et al., 2009; Simo & Bies, 2007). Moreover, organizations might need to adopt new organizational practices and routines to comply with data sharing agreements. Because of this, formal coordination might decrease autonomy while increasing interdependence among organizations (Azad and Wiggins, 1995; Malatesta & Smith, 2014; Weitzman et al., 2006). Finally, while formal coordination simplifies and speeds up the process of exchanging data by establishing routines, organizations might still encounter barriers if they want to access different data or modify the agreements to face emerging challenges or needs. Formal coordination is limited to what organizations have agreed upon (Willem & Buelens, 2007).

I suggest that managers who utilize formal coordination mechanisms are more likely to access data from other public agencies. Previous studies show that formal coordination is common in inter-organizational relationships generally (Bryson et al., 2006), and specifically in inter-organizational relationships among public agencies (Allard et al. 2007; Weitzman et al., 2006). The vertical structure of the government bureaucracy influences relationships among public agencies and promotes values such as authority, clarity of roles and responsibilities, and accountability (Agranoff & McGuire, 2001; Eglene, Dawes, & Schneider, 2007; Fountain, 2007). Formal coordination is well suited under these conditions.

In public organizations, a shared bureaucratic culture increases social integration, defined as "the degree to which ind-ividuals are institutionally bound and accountable to a bounded group or collectivity" (6, 2007, p. 408-409). Public organizations and employees are subjected to the same institutional structure and respond to similar professional and cultural norms (Fountain, 2007). For instance, public managers respond to higher accountability standards than nonpublic managers and can rely on similar professional training to provide guidance on acceptable behaviors and expectations (Kornberger, Meyer, Brandtner, & Höllerer, 2017). Shared norms and culture increase trust and cognitive homogeneity among public employees, which decrease the costs and time needed to develop shared agreements (Nahapiet and Ghoshal, 1998).

At the same time, relationships among public organizations are characterized by clear divisions across government levels, branches and jurisdictional areas (Agranoff & McGuire, 2001; Daley, 2009; Dawes et al., 2009). These divisions can hinder public managers' ability and willingness to work together and exchange data and information. Public managers might compete against each other or pursue incompatible missions, which increases the likelihood of conflict. Moreover, the diversity of rules and regulations that bind the actions of different public agencies might slow down collaboration and interaction (Daley, 2009).

Formal coordination is adapted to access data from other public agencies because formal agreements and routines align well with the bureaucratic values of authority and accountability that characterize public sector relationships. Formal coordination promotes organization accountability by clearly assigning responsibilities, roles, and costs, and ensures that organizations comply with rules and regulations that govern their functioning. Moreover, formal coordination addresses issues of jurisdictional competence by establishing limits and purposes of data access. For instance, agreements can establish norms limiting the use of data for specific purposes or the transfer of data to a third party (Susha, Janssen, & Verhulst, 2017). Finally, the underlying common culture and norms across public organizations facilitate the negotiations of formal agreements because public employees act within a common cognitive, regulatory, and cultural framework. Therefore, I suggest that:

H3a: Formal coordination mechanisms will be positively correlated with data access from other public agencies.

Lateral coordination. Lateral coordination occurs when managers submit a formal data request to another organization or develop data sharing agreements that regulate occasional data transfers. Compared to formal coordination, organizations that engage in lateral coordination do not develop routines for sharing data. Instead, data sharing remains unplanned and subject to case-by-case agreements. When public managers submit a formal request, the receiving organization can decide whether to accept the request or deny it and not provide data.

There are potential advantages and disadvantages to lateral coordination. Because it is unplanned, lateral coordination allows organizations to manage privacy and conflicts on a case-by-case basis, thereby maintaining control over data use and transfer (Willem & Buelens, 2007). Organizations can also learn from their experience and adjust over time depending on the outcomes of previous exchanges (Willem & Buelens, 2007). Furthermore, lateral coordination allows organizations to explore new opportunities to exchange data with a variety of actors. Relationships are less formalized and stable and negotiation time and costs are low because the agreements can be re-negotiated and are not binding in the long term (Galbraith, 1973).

However, lateral coordination increases the risk of not obtaining access to data because organizations can refuse to fulfill the request. Cultural, organizational, and institutional barriers across organizations might hinder data sharing because lateral coordination does not adequately address concerns about data misuses and security. Moreover, public managers might need to engage in lengthy negotiations and await an uncertain outcome. The length of the process might negatively affect the timeliness of data access and therefore the relevance of the information that public managers can extract from data.

I suggest that lateral coordination is positively correlated with data access in cross-sector relationships. Because of its flexibility, lateral coordination is a good fit for complex contexts where formal coordination can be a costly and lengthy process (Galbraith, 1973; Willem & Buelens, 2007). It is also more suitable when organizations are less dependent on one another and do not wish to sustain the costs of developing shared routines or adopting interoperable systems (Fishenden & Thompson, 2013; Willem & Buelens, 2007). All these conditions apply in cross-sector relationships.

Cross-sector relationships are increasingly common in the public sector (Bryson et al., 2006; Meerkerk & Edelenbos, 2017; Quick & Feldman, 2014). However, while public agencies are often legally required to interact with one another, interactions with nongovernmental organizations occur at the initiative of individual organizations. Cross-sector relationships are complicated by stark cultural and cognitive differences related to the diversity of norms and values between public, nonprofit, and private organizations which increase conflicts and hinder collaborative agreements (Bryson et al., 2006; Daley, 2009; Guo & Acar, 2005). For instance, private organizations have lower accountability and transparency requirements and are often less willing to share data with other organizations. Reasons vary and might include fear of data misuse or concerns about competition (Graber, 1992). As nongovernmental organizations are less dependent on public agencies, they also have fewer incentives to collaborate, especially when collaboration requires them to adjust their own routines (Roberts, 2011).

Lateral coordination might facilitate data access in cross-sector relationships because it allows organizations to negotiate conditions for data access and change them over time. In lateral coordination, organizations do not engage in the long-term but address concerns that are specific to each data exchange. This flexibility is important when there are high cultural and cognitive barriers and organizations learn over time how to moderate them (Chen & Lee, 2018). Moreover, lateral coordination requires lower incentives for organizations to collaborate and does not reduce an organization's autonomy. Private for-profit and nonprofit organizations are more willing to cooperate under these conditions which allow them to engage with the public sector without losing their independence and by maintaining their routines and structure. Therefore, I expect that:

H3b: Lateral coordination will be positively correlated with data access from non-governmental organizations.

Informal coordination. Informal coordination occurs when public managers access data through personal relationships. Social exchange and social capital theory have long recognized the importance of social relationships in exchanging resources (Blau, 1955; Molm, Whitham, & Melamed, 2012; Nahapiet & Ghoshal, 1998; Tsai, 2002). Personal networks are linked to high trust and reciprocity, which decrease concerns about security and data misuse as well as cultural and cognitive barriers, thereby facilitating exchange and negotiation (Chen & Lee, 2018; Gil-Garcia, Pardo, & Burke, 2010). However, social networks also require time to be created and place high demands on public managers who need to invest additional efforts and resources to build personal relationships (Allard et al., 2018; Berman, West, & Richter, 2002). Social networks are based on social obligations and have high maintenance costs as managers need to maintain frequent interaction, reciprocate exchanges of resources and feed trust over time (Burt, 2000; Nahapiet & Ghoshal, 1998). Finally, interpersonal networks rely on individuals' willingness to provide data. Compared to both formal and lateral coordination, informal coordination might occur without written agreements or communication, and managers and employees have higher discretion in complying with data requests. The lack of formal or written agreements might also increase threats related to data use, security, and management of sensitive information.

Public managers have extensive personal and professional networks that span across public, private, and nonprofit organizations, both inside and outside the city boundaries. Public managers often leverage their networks to gain and control access to resources, including data and information (Aiken, Bacharach, & French, 1980; Malone & Crowston, 1994; Romzek & Dubnick, 1987). I suggest that informal coordination is positively correlated with data access in intra-organizational and cross-sector relationships.

Departmentalization is a natural barrier to the access of data and information across organizational units. Departmentalization refers to differences in the goals, values, culture, and structure of departments within an organization, which prevent public managers from drawing from other departments' experience and information (Argote, Gruenfeld, Naquin, & Turner, 2001; Daley, 2009; Drew, 2015). For instance, departments in different policy areas respond to different professional and social norms, which increase fragmentation within a city organization and decrease willingness to collaborate and share data (Cohen, 2017; Roberts, 2011).

Managers are likely to develop personal relationships within their organizations because the geographic proximity of members facilitates interaction (Tsai & Ghoshal, 1998). Additionally, intra-organizational relationships are characterized by a common sense of solidarity and a shared understanding of professional norms and organizational culture (Bolino et al., 2002; Romzek et al., 2014; Tsai & Ghoshal, 1998). These conditions facilitate data access because public managers can anticipate the behavior of their colleagues and are knowledgeable about limits, boundaries, and issues involved in data sharing (6, Bellamy, Raab, Warren, & Heeney, 2007). Finally, through personal relationships, public managers can ask questions about the data, metadata and data collection, and gain a better understanding of the value that data can provide. Communication reduces cognitive barriers and increases the likelihood of exchanging data (Inkpen and Tsang, 2005; Nahapiet and Ghoshal, 1998).

Informal coordination might also be effective in cross-sector relationships, where self-interest and diverse professional frameworks are the main barriers to data exchange (Dawes, 1996; Gil-Garcia et al., 2010). Previous studies show that conflicts are higher in cross-sector relationships (Simo & Bies, 2007), where actors are more likely to have different scopes and purposes for sharing data (6 et al., 2007). Preexisting relationships, both personal and professional, might generate trust and social norms that positively affect the likelihood of accessing data (Dawes et al. 2009). Previous research has used the term “boundary spanners” to indicate individuals that promote interaction across sectors with the scope of facilitating the transfer of information and other resources (Meerkerk & Edelenbos, 2017). Boundary spanners bridge and reconcile cognitive barriers that might hinder exchange (Aiken et al., 1980) and are associated with more effective management of conflicts that arise in cross-sector relationships (6 et al., 2007; Simo & Bies, 2007). I, therefore, expect that:

H3c: Informal coordination mechanisms will be positively correlated with data access from other departments in the city.

H3d: Informal coordination mechanisms will be positively correlated with data access from non-governmental organizations.

CHAPTER 5

DATA AND METHOD

This chapter describes data collection, variables, and method. The study focuses on a nationally representative sample of 500 US cities with populations from 25,000 to 250,000. In figure 4 a black dot indicates each city included in the study. Small- and medium-sized cities are a relevant research setting because they represent approximately 27% of the US population – more than 85 million individuals (US Census Bureau, 2010).

Data are drawn from a variety of sources. The survey data comes from the 2016 National Study of Technology in City Government titled “Data sharing, civic engagement, and technology us in local government agencies” and conducted by the Center for Science, Technology and Environmental Policy Studies (CSTEPS) at Arizona State University (ASU). The survey was used to collect department-level data about how often department heads request data; how often they obtain data, and from which other organizations; which barriers they encounter when requesting data; and coordination mechanisms they utilize more often. Other sources include data collected by the Sunlight Foundation on state open data portal and data provided by the Pew Charitable Trusts on state government institutional capacity.

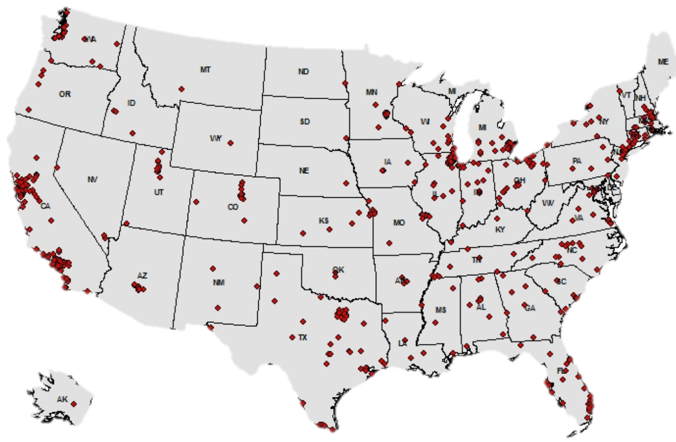

Figure 4. Map of 500 cities included in the study

Survey Data

The survey data were drawn from the National Study of Technology in City Government titled “Data sharing, civic engagement, and technology us in local government agencies” and conducted by the CSTEPS in 2016. The survey instruments consist of several questions about participation (e.g., frequency, stakeholder types, and legal requirements); technology use (e.g., purpose and frequency of technology use, technology management, perceived technology outcomes, and perceptions of technology); department climate and work environment (e.g., organization values and work-life balance); and data sharing (e.g., frequency of data exchange, barriers to data exchange, and coordination mechanisms).

The sample frame includes all 184 cities with populations ranging from 100,000 to 250,000 and a randomly selected sample of 316 cities with populations ranging from 25,000 to 99,999. The director or deputy director from five departments in each city were invited to participate: Mayor’s office, finance, police, community development, and parks and recreation. The sample includes a total of 2,500 department heads in 500 US cities.

The survey items were tested in previous versions of the survey that were conducted by CSTEPS in 2010, 2012, and 2014 on the same sample frame. The research team revised the items according to inputs from previous surveys. The data sharing section was newly designed to collect data for this study. I conducted a set of cognitive interviews (N=3) with public managers to ensure that the survey items did not include jargon or terms that public managers might find unclear or ambiguous. The interview protocol is presented in Appendix A. The interviews were conducted with managers from three departments - finance, Mayor’s office, and police[6] - in three cities whose size is between 25,000 and 250,000 inhabitants. The interviews asked questions to confirm that respondents were able to: understand the meaning of the survey items; provide the information that the instrument intended to collect; and find the best answer among the options provided. Results showed that the respondents correctly described the object of the inquiry and understand the term "data" in line with the definition provided in the survey. Respondents reported that questions were clear and understandable and did not report difficulties in selecting a response among the options provided. For instance, respondents confirmed that the data frequency categories (annually – quarterly – monthly – weekly) reflect a common terminology used in government activities. In some cases, respondents reported that the categories were redundant (e.g., the list of technological

---

[6] Current human resources manager previously employed in police departments. Questions focused on activities in the police department.

barriers) or that extreme options were not available (e.g., “my organization does not request data”). These items were modified or added to the survey as appropriate.

Participants’ contact information were collected from city websites or by calling the city and requesting contact information. The survey was administered online over three months, from September 27th to December 27th, 2016. After removing wrong and bad email addresses and managers who had retired or left their position, the sample was reduced to 2,166 eligible individuals. We received a total of 667 responses for a response rate of 30%, including both complete and partial surveys retained because the respondent answered more than half of the questions. The response rate was calculated following the procedure of the American Association for Public Opinion Researchers (RR2 - AAPOR, 2016).

We received responses from 385 cities from 45 states[7]. The average city size is 87,400 inhabitants. When comparing respondents to non-respondents, the research team found that respondents are more likely from community development, parks and recreation, and police departments ($p<0.05$) and they more frequently work for council-manager cities (75%). There is no significant difference in city size between respondents and non-respondents ($p>0.05$).

The data sharing section opened with a screening question asking respondents whether their department “obtains data generated by other organizations to do its work.” Only respondents who replied “Yes” were asked to respond to other questions concerning

[7] Our survey did not include responses from Hawaii, Maine, Montana, Rhode Island, and Vermont.

data sharing. Nearly two-thirds (70%) of the city department heads reported obtaining data from other organizations to do their work[8].

Other Data Sources

I combine survey data with other sources of data to measure the state-level institutional environment in which city departments are embedded. I utilize data from three sources: the Sunlight Foundation, the Pew Charitable Trusts, and Comparitech.

The Sunlight Foundation is “a national, nonpartisan, nonprofit organization that uses civic technologies, open data, policy analysis, and journalism to make our government and politics more accountable and transparent to all.” (Sunlight Foundation, n.d.). As part of its effort, the foundation has developed an "Open States" report that compares open data portals across state governments. Data were collected and made available in 2013. The research team integrated online data collected online with short interviews with state legislators to ensure the accuracy of the information collected. Changes to the original dataset are regularly reported in the methodology section on the Sunlight Foundation website[9].

The Pew Charitable Trusts conducted a study in 2018 entitled “How States Use Data to Inform Decisions. A national review of the use of administrative data to improve state decision-making” (Pew Charitable Trusts, 2018). The study focuses on administrative data, defined as “data collected and maintained by a federal, state, or local government; government agency; or contractor or grantee of the agency, primarily for the

[8] A discussion about possible selection biases in the analysis is provided in Appendix B. Estimation of an Heckman selection model shows no evidence of a selection bias.

[9] Full information can be found here: https://openstates.org/reportcard/#changelog

routine management of programs such as TANF, Medicaid, the corrections system, unemployment insurance systems, and child support payment systems" (p. 46). Pew Charitable Trusts undertook two data collection efforts. First, data were collected using the Lexis Advance database to identify "bills, statutes, and executive orders related to data use in eight major subject areas: governance, privacy, warehouse, inventory, integration, sharing, security, and integrity" (Pew Research Trusts, 2018, p. 46). Then, the research team interviewed 341 officials across all the 50 US states and the District of Columbia to understand data usage and institutions supporting data sharing. Interviews were conducted with officials in six government functions: auditing or evaluating, budgeting, performance management, legislative research, information technology management, and centralized data analytics or data management and human services agency officials. Interviews were analyzed using an iterative qualitative coding procedure to identify common topics and patterns across interviews. This research utilizes data collected through the search in the Lexis Advance database to measure state government institutional capacity. Data were made available in their original format by the Pew Charitable Trusts research team.

Finally, I utilize data collected by a journalistic initiative comparing state legislation on issues related to online privacy across US states[10]. The research compares state laws regulating how private and public organizations can disclose data and information and transfer them to third parties (Bischoff, 2017). To confirm the validity of

[10] Further information can be found here https://www.comparitech.com/blog/vpn-privacy/which-us-states-best-protect-online-privacy/

the research, I compared the results with other similar sources; results are similar with states that have stringent laws (e.g., California) and more “relaxed” states (Leuan, 2017).

Variables

This section describes the measurement and descriptive statistics of the variables employed in the models that will be discussed in Chapter 6.

Dependent variable: data access. In the survey, we asked managers to think about “data that your organization uses for its activities such as organizational performance, employee behavior, transactions, citizen, businesses or other non-profit activity, budget and financial statistics, geospatial data (e.g., GIS data), and so on." We then asked: “Approximately what share of your organization's requests for data are fulfilled without requiring your organization to follow up or make additional requests?”. The question aims to capture the effectiveness of department heads to access data from other organizations in a timely manner. The question was asked for three types of stakeholders:

(1) "Other governmental departments in your city or town"

(2) "Government organizations outside your city (other cities, county, state, federal)"

(3) "Non-governmental organizations (private and non-profit)"

Respondents could indicate one of the following response categories: “No requests”, “Few requests”, “Some requests”, “Most requests”. The survey item is depicted in figure 5.

	No requests	Few requests	Some requests	Most requests
Other governmental departments in your city or town	☐	☐	☐	☐
Government organizations outside your city	☐	☐	☐	☐
Non-governmental organizations (private and non-profit)	☐	☐	☐	☐

Figure 5. Survey items used to construct the dependent variable

From the survey question, I develop three categorical dependent variables labeled: Data Access from Other Departments in the City; Data Access from Other Public Agencies; and Data Access from Non-Governmental Organizations. Given the few observations in the "No Request" category (respectively, 2.7%; 3.4%; 8.3%), which might cause problems in the model estimation, I grouped the categories "No Requests" and "Few Requests" into one category. Therefore, each variable includes three categories: 1 = "None to few requests - Low data access"; 2 = "Some requests - Medium data access"; and 3 = "Most requests - High data access".

Figure 6 reports the frequency of respondents for each category of each dependent variable. On average, city department heads report greater data access from other departments in the city, with 55% of respondents indicating that "most requests" are fulfilled without the need to follow up; it follows data access from other public agencies (45%) and non-governmental organizations (37%).

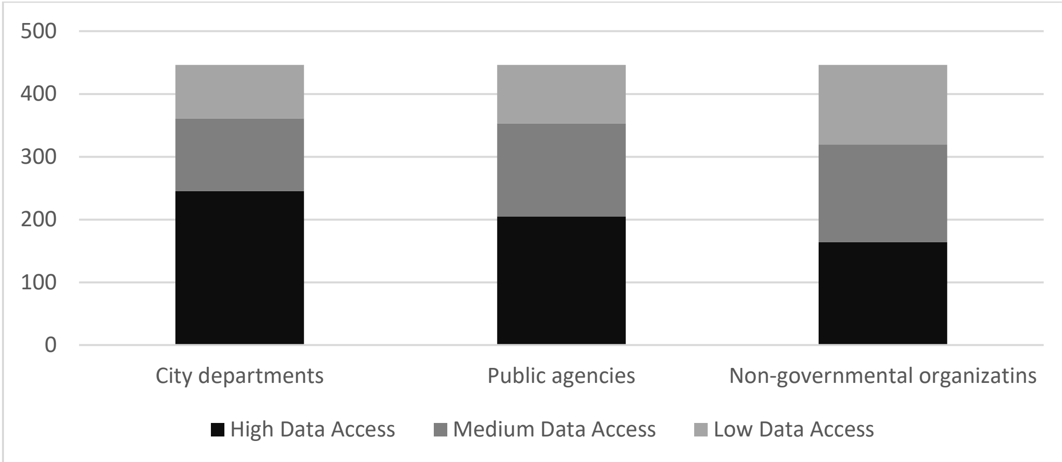

Figure 6. Data access, frequency by stakeholder type

Institutions. Institutions include four variables: Privacy Laws; Legal Mandate; Institutional Capacity; and Open Data Portals.

Privacy Laws is a continuous variable ranging from 0 to 100, which serves as a proxy for a state government's engagement and attention to the development and enactment of privacy laws. The score is based on the fourteen criteria reported in table 2. The average score is 60.01 (s.d. 16.06). The highest score is 85.70 in California and Delaware. The lowest score is 28.6 in Alabama, South Dakota, and Wyoming.

Table 2.

Number of states implementing privacy laws, by privacy law type.

Criteria	# of states
1. Internet Service Providers barred from sharing info with third parties	2
2. Internet Service Provides require explicit consent to share customer data	50
3. Must dispose of customer data after set period of time (Government)	15
4. Must dispose of customer data after set period of time (Companies)	31
5. Require to disclose when a breach occurs (Companies	47
6. Laws protect privacy of e-reader data (Libraries)	4
7. Social media privacy (employers	26
8. Social media privacy (educational institutions)	16
9. Shield law to protect journalists (SL)	40
10. Shield law to protect journalists (CRP)	36
11. Laws to protect employee privacy	5
12. Laws to protect K-12 student information	36
13. Laws to protect children's privacy	50
14. Laws to protect Internet of Thing data.	1

Source: Compari Tech. Retrieved 3rd March 2018 from https://www.comparitech.com/blog/vpn-privacy/which-us-states-best-protect-online-privacy/

Legal Mandate is a department-level variable measuring the extent to which external organizations are mandated to provide data to the city department. Respondents could indicate if other departments in the city, other public agencies, and non-governmental organizations were legally requested to provide “all”, “most”, “some”, or “no” data to them. I create three continuous independent variables: Legal Mandate - Other Departments in the City, Legal Mandate - Other Public Agencies, and Legal Mandate - Nongovernmental organizations. All variables range from 1 = “no legal requirements” to 4 = “all requests are subject to legal mandate”. The average of Legal Mandate - Other Departments in the City is 2.45 (s.d. 1.08). The average for Legal

Mandate - Other Public Agencies is 2.37 (s.d. 0.97). The average for Legal Mandate - Nongovernmental organizations is 1.66 (s.d. 0.82).

Institutional Capacity is a state-level variable based on the Pew Charitable Trusts data on data sharing laws across US states (Pew Charitable Trusts, 2018). The variable counts the number of statewide laws concerning governance, warehouse, inventory, integration, sharing, agreement, security, and integrity of data across all policy areas. Detailed information per each state is reported in Appendix C. The variable ranges from 0 to 10, with an average of 4.95 (s.d. 3.41).

Open Data Portal is a state-level categorical variable that indicates the quality of open data portals across state governments. For each state, researchers from the Sunlight Foundation evaluated six dimensions of open data quality: completeness, timeliness, ease of access, machine readability, use of common standards, and permanence. The criteria used to evaluate each dimension are reported in Appendix D. The maximum score is 7 while the mean is 2.65 (s.d. 1.24). Alabama, Kentucky, Massachusetts, and Nebraska scored the highest. Several states received a minimum score of 1 point: Arkansas, Connecticut, Georgia, Kansas, North Carolina, New Hampshire, New York, Pennsylvania, Texas, Virginia, and Washington.

Social environment. The social environment is measured using three variables derived from the survey data: City Actor Influence, Government Influence, and Civil Society Influence. The survey asked respondents to indicate how influential a set of actors are on the city's policy-making process. Responses ranged from 1 = "No influence" to 5 = "Strong influence". Actors include the mayor, mayor's council, other city departments, governor, state legislature, state courts, the federal government, business groups, advocacy groups, public opinion, and media. A factor analysis shows that three factors can be extracted (eigenvalues > 1) as shown in table 3.

Table 3.

Influence items, factor analysis.

	Component		
	1	**2**	**3**
Mayor	0.10	0.10	0.87
Mayor's Council	0.11	0.16	0.84
Other city departments	0.14	0.34	0.56
Governor	0.78	0.08	0.18
State legislature	0.82	0.11	0.14
State courts	0.81	0.12	-0.03
Federal governments	0.72	0.15	0.12
Business groups	0.26	0.72	0.22
Advocacy groups	0.15	0.79	0.10
Public opinion	-0.03	0.80	0.12
Media	0.14	0.70	0.17

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 4 iterations.

The three factors reflect (1) influence from city actors, such as the mayor, mayor’s council and other city departments, (2) influence from higher government levels, such as the governor, state legislature, state courts and the federal government, and (3) influence from the civil society, including business and advocacy groups, public opinion and media. I created three average scales ranging from 1 to 5 for each group of actors. City Actor Influence has a Cronbach’s alpha equal to 0.73 and a mean of 3.53 (s.d. 0.86). Government Influence has a Cronbach’s alpha equal to 0.84 and a mean of 2.25 (s.d.

0.79). Civil society Influence has a Cronbach's alpha equal to 0.81 and a mean of 2.62 (s.d. 0.74).

Coordination mechanisms. I measure four types of coordination mechanisms: Formal Coordination; Lateral Coordination; and Informal Coordination. Formal coordination is measured by asking respondents: "My organization has well established routines to regularly receive data from other organizations". The response scale ranges from 1 = "Strongly disagree" to 5 = "Strongly agree." Average is 3.7 (s.d. = 0.79). Lateral coordination measures how frequently the city department submits formal written or online requests to other organizations to obtain data. Scale ranges from 1 = "Never" to 5 = "Always"; the average is 2.95 (s.d. = 0.82). Informal coordination measures how often a department head contacts someone he or she knows in another organization to get access to the data. Scale ranges from 1 = "Never" to 5 = "Always". The average is 3.62 (s.d. = 0.82).

Table 4 shows the correlation between each pair of coordination mechanisms. Lateral coordination is most strongly correlated with other coordination mechanisms; the correlation between formal and lateral coordination is equal to 0.14 and between informal and lateral coordination is 0.16. The low correlations across coordination mechanisms suggest that respondents perceive each coordination mechanism as a distinctive strategy to access data.

Table 4.

Correlation across coordination mechanisms.

	Formal	**Lateral**	**Informal**
Formal	1.00		
Lateral	0.14	1.00	
Informal	0.02	0.16	1.00

Control variables. Control variables include Technological Barriers, Socio-political Barriers, Technical Capacity, Department Type, Principal City, Population Size, Department Size, and Form of Government.

To account for barriers that might prevent access to data, I include two variables, Technological Barriers and Socio-political Barriers. Technological barriers are widely cited in the literature and refer to situations where “data were not stored electronically” or “data were not transferable because of incompatibility across information systems” (Dawes, 1996). Socio-political barriers include items such as “data were too politically sensitive to be shared” or “management did not want to share data because of fear of public criticism”. Each item ranges from 1= "Never" to 5 = "Very Often". Table 5 shows the full list of items.

I performed a factor analysis to confirm the two scales. The results are presented in Table 5. The factor analysis shows two factors, one including technical barriers and one including socio-political elements. Each factor has an eigenvalue greater than 1 and the total variance explained is 58.5%. I created two scales by averaging the items. The Cronbach’s alpha is 0.70 for the Technological Barriers scale and 0.82 for the Socio-political Barriers scale, showing good and excellent fit, respectively. The technical barriers scale has an average of 2.78 (s.d. = 0.71), while the socio-political barriers scale has an average of 2.28 (s.d. = 0.7) indicating that respondents are more likely to encounter technical rather than socio-political barriers.

Table 5.

Technical and Socio-political barriers to data exchange.

Rotated Component Matrix		
Variables	**Component**	
	Socio-political	**Technical**
The other organization did not have the requested data	-.125	.724
The requested data was not electronically stored or available in a retrievable electronic format.	.064	.787
The data were not transferable because of incompatibility across information systems.	.311	.716
Our organization was not equipped to store, receive, or analyze the data.	.406	.515
Because of regulatory and privacy issues, the other organization was prohibited from sending us the data.	.611	.381
There were too many rules and levels of approval to access the data (i.e. written consent, legal authorization, court orders, etc.).	.672	.336
The data were too politically sensitive to be shared.	.772	.111
The management did not want to share the data because of fear of public criticism.	.870	-.052
The management did not want to share the data because of competing interests with our organization	.782	-.012

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 3 iterations.

Technical Capacity is an average scale of five items: “My agency is ill-equipped to manage important questions about online security and privacy”, “Staff in my office are resistant to change related to technology”, “Management lacks software applications that would make work more efficient", "There is a mismatch between our department's needs and what technology can provide", and "My agency is too busy to effectively monitor, control, and use the data we collect". A factor analysis shows that items load on one factor; a Cronbach's alpha of 0.73 confirms the reliability of the scale. The average for the scale is equal to 3.28 (s.d. = 0.68).

File sharing technology measures whether a city department utilizes information sharing tools such as cloud-based technologies (e.g. Dropbox, Google drive). The variable is coded as 1 if the department utilizes file sharing technology or 0 if it does not. On average 79% of departments utilize data sharing technologies.

Department Type is a set of dummy variables indicating the department type. The reference category is the Mayor's Office. Among our respondents, 14.5% work in the mayor's office, 28% in community development, 16.5% in finance, 19% in parks and recreation, and 22% in the police department. Previous research has found that variation in data sharing practices is a function of the policy sector (6, Bellamy, Raab, Warren, & Heeney, 2007) and suggested that departments might be subject to different professional and normative standards (Grimmelikhuijsen & Feeney, 2016).

Department Size is the self-reported number of full-time employees (FTEs) working for the department. Larger departments might have more resources and a greater influence on their external environment. On average, department size is 158.63 FTEs, but there is great variation in the sample with a standard deviation of 425.63 FTEs. The variable is logged to account for this variation. I also include a Principal City variable and a Population variable. Principal City is a dummy variable that indicates whether the department is a principal center of a metropolitan area (=1) or not (=0). About one third (27%) of the cities in the sample is a principal city. The Population variable is included in logarithmic form (mean = 11.2, s.d. = 0.66) and draws from the population reported by the 2005 US Census Bureau. Finally, Form of Government indicates whether the city is a Mayor-Council (= 1) or a Council – Manager (= 0) government. 24% of the cities in the sample are a Mayor-Council government.

Descriptive statistics are reported in table 6, while correlations are reported in Appendix E.

Table 6.

Descriptive statistics.

	Obs.	Mean	St. Dev.	Min	Max
Privacy Laws	463	60.00	16.10	28.60	85.70
Legal Mandate-Other depts. in the city	447	2.40	1.10	1.00	4.00
Legal Mandate-Other public agencies	445	2.40	1.00	1.00	4.00
Legal Mandate-Non-govern. Org.	447	1.70	0.80	1.00	4.00
Institutional Capacity	463	4.95	3.42	0.00	10.00
Open Data Portals	463	2.65	1.24	1.00	5.00
City Influence	463	3.53	0.86	1.00	5.00
Government Influence	463	2.25	2.25	1.00	5.00
Civil Society Influence	463	2.62	2.62	1.00	5.00
Formal Coordination	458	3.70	0.79	1.00	5.00
Lateral Coordination	439	3.05	0.82	1.00	5.00
Informal Coordination	440	3.62	0.82	1.00	5.00
Technological Barriers	437	2.78	0.71	1.00	5.00
Socio-political Barriers	435	2.28	0.70	1.00	4.80
Technical Capacity	462	3.28	0.68	1.00	5.00
File Sharing Technology	444	0.79	0.41	0.00	1.00
Principal City	463	0.27	0.45	0.00	1.00
Mayor's office	463	0.14	0.35	0.00	1.00
Community development	463	0.28	0.45	0.00	1.00
Finance	463	0.16	0.37	0.00	1.00
Parks and recreation	463	0.19	0.39	0.00	1.00
Police	471	0.22	0.41	0.00	1.00
Population Size (log)	463	11.2	0.66	10.14	12.43
Department size	447	158.63	425.63	1.00	4500.00
Department size (log)	447	3.71	1.60	0.00	8.41
Form of Government - Mayor Council	463	0.24	0.43	0.00	1.00

CHAPTER 6

DATA ANALYSIS

This chapter discusses the data analysis, which investigates variation in data access across the portfolio of relationships of a city department: other departments in the same city; other public agencies; and non-governmental organizations. The chapter is divided into two sections. The first section discusses the model fit and presents the estimation of three models: an ordered logit model with clustered standard errors, a multilevel ordered logit model, and a seemingly unrelated regressions model. All models were estimated using Stata v.15, including user-written command cmp (Roodman, 2011) for the seemingly unrelated regressions model and gologit2 for the generalized ordered probit models (Williams, 2006). Despite some minor differences, the three models provide consistent results, both considering the significance and sign of the coefficients. The second section discusses the results, including statistical interpretation of the coefficients and hypotheses testing.

Model Estimation

The analysis focuses on three dependent variables: Data Access from Other Departments in the City; Data Access from Public Agencies; Data Access from Non-Governmental Organizations. Each dependent variable (DV) contains three ordered categories: "low data access", "medium data access" and "high data access". Therefore, I utilize an ordered model, which is appropriate when the dependent variable consists of ordinal categories (Long, 1997). Ordered categorical variables are censored variables as respondents would have been able to identify their position along a continuous variable if given a choice (Long, 1997). The estimation of an ordered model assumes an underlying continuous variable y* observed through y. The variable y provides incomplete information because it is observed only within specified categories; the τ are thresholds or cut-points of each category, whereas the extreme categories are delimited by -∞ and +∞, so that:

$$y \begin{Bmatrix} 1 \rightarrow Low\ Access & if\ \tau = -\infty \ \leq \ y_i^* < \ \tau_1 \\ 2 \rightarrow Medium\ Access & if\ \tau_1 \ \leq \ y_1^* \ < \tau_2 \\ 3 \rightarrow High\ Access & if\ \tau_2 \ \leq \ y_i^* \ < \ \tau_3 = \ +\infty \end{Bmatrix}$$

The underlying variable y* is used to fit the vector β.

Multicollinearity. Before estimating the models, I checked for multicollinearity issues. Multicollinearity occurs when one independent variable can be linearly predicted by other independent variables with a fair degree of accuracy (Lewis-Beck & Lewis-Beck, 1980). Multicollinearity might cause inflated standard errors, model sensitivity to small changes, and incorrect estimation of the parameter sign (Greene, 2000). Since multicollinearity is a property of the β vector - not of the model - it can be tested before estimating the model.

The variance inflation factor (VIF) is used to check for multicollinearity; a high VIF indicates that a variable is not orthogonal to the other variables, thereby implying collinearity. Researchers consider worrisome a VIF higher than ten; it corresponds to a tolerance factor lower than 0.1. The square root of the VIF indicates how much larger each standard error is, as compared to its value in a model in which variables are completely uncorrelated. Results of the multicollinearity analysis are reported in table 7. The analysis accounts for missing values in each dependent variable. None of the variables has a square root VIF higher than 2, suggesting that there is very low multicollinearity in the data (Lewis-Beck & Lewis-Beck, 1980). The average VIF in each model is equal to 1.56. Overall, the data do not present multicollinearity issues.

Table 7.

VIF estimation.

	Data access from other depts. in the city		**Data access from other public agencies**		**Data access from non-gov. orgs.**	
	VIF	VIF^2	VIF	VIF^2	VIF	VIF^2
Institutions						
Privacy Laws	1.75	1.32	1.73	1.32	1.74	1.32
Legal Mandate	1.11	1.05	1.14	1.07	1.10	1.05
Instit. capacity	1.95	1.40	1.94	1.39	1.95	1.40
Open govern.	1.35	1.16	1.37	1.17	1.36	1.17
Social environment						
City Influence	1.53	1.24	1.52	1.23	1.52	1.23
Civil Soc. Influence	1.68	1.29	1.67	1.29	1.66	1.29
Government Influence	1.51	1.23	1.50	1.22	1.50	1.23
Coordination mechanisms						
Formal Coordination	1.18	1.09	1.19	1.09	1.17	1.08
Lateral Coordination	1.15	1.07	1.14	1.07	1.17	1.08
Informal Coordination	1.12	1.06	1.11	1.05	1.12	1.06
Control variables						
Principal City	1.31	1.14	1.28	1.13	1.29	1.14
Social Barriers	1.26	1.12	1.27	1.12	1.26	1.12
Technical Barriers	1.25	1.12	1.25	1.12	1.26	1.12

Dept. Size	1.95	1.40	1.93	1.39	1.94	1.39
Technical Capacity	1.15	1.07	1.16	1.08	1.15	1.07
File Sharing Tech.	1.14	1.07	1.14	1.07	1.13	1.07
Form of Government	1.25	1.12	1.23	1.11	1.23	1.11
Population	1.62	1.27	1.61	1.27	1.62	1.27
Comm. Development	2.65	1.63	2.62	1.62	2.62	1.62
Finance	2.16	1.47	2.16	1.47	2.13	1.46
Parks and Recreation	1.98	1.41	1.98	1.41	1.98	1.41
Police	2.50	1.58	2.50	1.58	2.44	1.56
Mean VIF	**1.57**					

Weights. All analyses are weighted to reflect the sampling procedure. As mentioned in Chapter 5, the sample was designed to include all 184 cities with populations ranging from 100,000 to 250,000 and a randomly selected sample of 316 cities with populations ranging from 25,000 to 99,999. Given the sampling strategy, proportional sampling weights are used to adjust for the probability that a city was included in the sample. The proportional sampling weights are the inverse of the probability to be included in the sample.

Table 8 reports the design weights; they were calculated based on city size. Cities with populations between 100,000 and 250,000 have a weight equal to 1, while small cities are assigned a weight between 3.159 and 3.177. Weights are implemented using the ‘svy’ framework in Stata.

Table 8.

Weights applied in the analysis.

Weight factor	N cities pop	N cities sample	Weights (pop/sample)
25K-50K	591	186	3.1774
50K-75K	278	88	3.159
75K-100K	133	42	3.1667
100-125K	68	68	1
125-150K	37	37	1
150-175K	23	23	1
175-200K	28	28	1
200-225K	18	18	1
225-250K	10	10	1
Total	1186	500	

Ordered Logit Model with Clustered Standard Errors

As a first step, I estimated each model using an ordered logit estimator with clustered robust standard errors (clustered SE). Clustered SEs are appropriate because observations are nested within cities and states; therefore, the errors terms are not independent but correlated within clusters. Violations of the assumption of independence of the error terms cause overestimation of the statistical significance even when the interclass correlation is small (ICC) (Cameron, Gelbach, & Miller, 2012; White, 1980). When data are clustered within different levels (city - state), researchers suggest clustering at the highest level, as computation of the SEs will be more conservative (Cameron & Miller, 2015). Clustered standard errors also adjust for heteroscedasticity issues, and they are referred to as clustered robust standard errors. I present the equation of each regression below:

$$\begin{aligned}
&\boldsymbol{Data\ access} = \\
&\beta_0 + \beta_1 Privacy\ Laws + \beta_2\ Legal\ Mandate + \beta_3 Institutional\ capacity \\
&\quad + \beta_4\ Open\ Data\ Portals + \beta_5\ City\ Influence + \beta_6 Civil\ Society\ Influence \\
&\quad + \beta_7\ Government\ Influence + \beta_8\ Formal\ Coordination \\
&\quad + \beta_9\ Lateral\ Coordination + \beta_{10}\ Informal\ Coordination \\
&\quad + \beta_{11} File\ Sharing\ Technology + \beta_{12} Technical\ barriers + \beta_{13} Socio \\
&\quad - political\ barriers + \beta_{14}\ Principal\ City + \beta_{15}\ Tech\ capacity \\
&\quad + \beta_{16}\ Form\ of\ government + \beta_{17}\ Population\ (log) \\
&\quad + \beta_{18} Department\ type + \beta_{19} Department\ size + \varepsilon_g
\end{aligned}$$

Where: $g = 1, \ldots, G$

The results of the three ordered logit models with clustered robust errors are reported in table 9. I reported coefficients, clustered standard errors, and the odds ratio transformation of the coefficients. Statistically significant coefficients (p-value < 0.1) are bold. All models contain 402 observations clustered in 44 states. The "data access from other departments in the city" model has a pseudo-R-squared equal to 0.083. The model is significant compared to a null model containing only control variables (Wald test, probability < 0.001 - chi2(11) = 53.10). The "data access from other public agencies" model has pseudo-R-squared equal to 0.08. The model is significant compared to a null model containing only control variables (Wald test, probability < 0.001 - chi2(11) = 46.59). The "data access from non-governmental organizations" model has a pseudo-R-squared equal to 0.075. The model is significant compared to null model containing only control variables (Wald test, probability < 0.001 - chi2(11) = 59.82). In all models the log likelihood increases, while the AIC and BIC decrease, as we move from a null to a full model, suggesting that the full model is an improvement of the null model. Table 9 reports odds ratio to simplify interpretation of the coefficients. An odds ratio shows an increase in the probability of an outcome when the independent variable increases of one unit.

Table 9.

Data access from other departments in the city, other public agencies, and non-governmental organization - Logit model with clustered robust SE.

	Data access from other departments in the city				**Data access from other public agencies**				**Data access from nongovernmental organizations**			
	Beta	**Odds ratio**	**SE**		**Beta**	**Odds ratio**	**SE**		**Beta**	**Odds ratio**	**SE**	
Institutions												
Privacy Laws	0.00	1.00	0.01		0.00	1.00	0.01		-0.01	0.99	0.01	
Legal Mandate	**0.38**	**1.46**	**0.16**	*****	**0.34**	**1.40**	**0.14**	*****	**0.26**	**1.30**	**0.13**	*
Institutional capacity	0.01	1.01	0.04		0.04	1.04	0.03		**0.07**	**1.07**	**0.04**	+
Open Data Portals	0.03	1.03	0.08		-0.13	0.88	0.09		-0.13	0.88	0.10	
Social environment												
City Influence	0.05	1.05	0.17		-0.03	0.97	0.12		0.01	1.01	0.12	
Civil Society Influence	**0.35**	**1.43**	**0.20**	**+**	**0.53**	**1.70**	**0.17**	******	**0.52**	**1.68**	**0.20**	**
Government Influence	-0.24	0.79	0.18		**-0.39**	**0.68**	**0.14**	******	**-0.37**	**0.69**	**0.14**	**
Coordination mechanisms												
Formal Coordination	0.09	1.09	0.16		0.23	1.26	0.15		0.15	1.16	0.18	
Lateral Coordination	-0.07	0.93	0.15		0.10	1.10	0.18		0.31	1.36	0.16	*
Informal Coordination	**0.43**	**1.53**	**0.14**	******	**0.33**	**1.40**	**0.17**	*****	0.21	1.23	0.18	
Control variables												
Principal City	**0.77**	**2.16**	**0.28**	******	0.48	1.62	0.30		**0.71**	**2.03**	**0.32**	*
Social Barriers	**-0.40**	**0.67**	**0.22**	**+**	**-0.45**	**0.64**	**0.21**	*****	**-0.31**	**0.73**	**0.16**	+
Technical Barriers	0.18	1.20	0.21		0.10	1.11	0.21		0.10	1.10	0.16	

File Sharing Technology	0.39	1.48	0.26		0.16	1.17	0.26		**0.68**	**1.98**	**0.25**	**
Dept. Size	-0.01	0.99	0.10		-0.09	0.91	0.10		-0.03	0.97	0.09	
Technical Capacity	-0.32	0.72	0.21		-0.22	0.80	0.17		**-0.38**	**0.68**	**0.15**	**
Form of Government	-0.15	0.86	0.24		-0.04	0.96	0.23		0.24	1.27	0.29	
Population	**-0.39**	**0.68**	**0.16**	*	-0.18	0.84	0.17		**-0.41**	**0.66**	**0.15**	**
Community Development	-0.56	0.57	0.40		**-0.77**	**0.46**	**0.39**	*	-0.25	0.78	0.39	
Finance	-0.64	0.53	0.53		-0.61	0.54	0.45		-0.20	0.82	0.37	
Parks and Recreation	-0.70	0.50	0.46		-0.63	0.53	0.46		-0.50	0.61	0.33	
Police	0.53	1.70	0.55		**0.84**	**2.31**	**0.43**	+	**0.75**	**2.11**	**0.36**	*
Cut 1	2.37	10.73	-8.57		-2.09	0.12	2.01		-3.97	0.02	1.61	
Cut 2	-2.46	0.09	2.32		-0.34	0.71	1.96		**-2.24**	**0.11**	**1.63**	***
Obs.		402				402				402		
Log pseudo-likelihood		-840.18				-898.48				-943.95		
Null model log likelihood		-914.68				-978.47				-1020.12		
Pseudo R2		0.08				0.08				0.07		
AIC		1728.36				1844.96				1935.89		
BIC		1824.27				1940.87				2031.81		

Significance levels: + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ - Reference categories: Mayor’s Office

Parallel line assumptions. Ordered models rely on the parallel line assumption stating that the cumulative probability curves are parallel for each category j. Because of this, the relationship between each category j and the independent variables is the same. In other words:

$$\Pr(Y = J_i|X) = \quad (x\beta_1)$$

When the assumption is true, we can assume that:

$$\beta_1 = \beta_2 = \beta_3$$

The Brant test (Brant, 1990) is usually employed to check the parallel line assumption in ordered models. It is a conservative test which often rejects the null hypothesis that the coefficients are equal, even if the assumption does hold (Peterson & Harrell, 1990). Moreover, when survey weights are applied, the Brant test is not appropriate; a Wald test is preferable.

The ologit2 package in Stata estimates a series of Wald tests on each variable to check if coefficients differ across categories and if the parallel line assumption holds. The test progressively constraints the variable with the least significant Wald test to have a coefficient equal across equations (i.e., to fit the parallel line assumption). Stata repeats the process until all variables for which the Wald test is non-significant are constrained to have equal coefficients across categories. When the process ends, if the Wald test remains not significant, the assumption is met.

In the "data access from other departments in the city" model, there is no evidence to reject the parallel line assumption. Instead, in the "data access from other public agencies" and "data access from non-governmental organizations" models, there is evidence that a few coefficients do not meet the parallel line assumption. In the "data access from other public agencies" model, the informal coordination variable (p-value <0.05) and the privacy variable (p-value < 0.01) do not meet the parallel line assumption. In the "data access from non-governmental organizations" model, the privacy variable (p-value<0.05) does not meet the parallel line assumption.

Researchers have different opinions about how to estimate the model when the parallel line assumption is not met. Researchers frequently switch to a multinomial model, where coefficients freely vary across categories. The downside of a multinomial approach is that we lose the information contained in the ordering when the model is estimated. Therefore, some researchers advise against switching to a multinomial model; the switch implies to move from a known but misspecified model to one that is merely suspect (Long, 1997). Finally, in a multinomial model, some parameters might become statistically insignificant because of the higher number of parameters that need to be estimated (Williams 2005).

Another option is to estimate a partial non-proportional odds model that relaxes the parallel line assumption only for the variables that failed the Wald test (Long & Freese, 2005). This option is helpful to observe how these variables behave when their coefficients are allowed to vary. I estimate the partial non-proportional odds model using the generalized ordered logit framework in Stata (gologit2) and selecting the coefficients that should be allowed to vary. The autofit option in the gologit2 package “uses an iterative process to identify the partial proportional odds model that best fits the data” (William, 2005, p. 3). Results for the variables that failed the parallel line assumption test are summarized in table 10. The interpretation of results for a partial non-proportional odds model does not differ from an ordered logit model.

In the "data access from other public agencies" model, informal coordination increases the likelihood to report low data access vs. the likelihood to report medium or high data access. The coefficient is positive and significant. By contrast, the coefficient is not significant when we compare low and medium data access vs. high access. Similarly, the privacy law coefficient is significant when comparing low vs. medium and high data access, but not when comparing low and medium vs. high data access.

In "data access from non-governmental organizations" model, the variable privacy law shows opposite sign when comparing low vs. medium and high data access and low and medium vs. high data access, but it is not significant.

Table 10.

Estimation of betas, standard errors, and p-values for variables that did not hold the parallel line assumption.

Public agencies	**B**	**SE**	**P-Value**
Informal coordination (1 vs 2+3)	.57	.22	0.01
Informal coordination (1+2 vs 3)	. 18	.17	0.28
Privacy (1 vs 2+3)	.02	.01	0.1
Privacy (1+2 vs 3)	-.009	.01	0.25
Non-governmental orgs.	**B**	**SE**	**P-Value**
Privacy (1 vs 2+3)	.003	.008	0.70
Privacy (1+2 vs 3)	-.013	.008	0.12

1 = low data access; 2 = medium data access; 3 = high data access

Multi-level Model

Given the data structure, a multilevel model could also be appropriate. A multilevel model assumes that factors at two or more levels explain the variation in the DV. In this case, the model includes variables at the department level (level-1) and variables at the state level (level-2). The equation for the model is specified as follows:

$$\begin{aligned}
&\boldsymbol{Data\ access_{ij}}(\boldsymbol{level\ 1}) = \\
&\beta_{0j} + \beta_{1j} Legal\ Mandate_{ij} + \beta_{4j} City\ Influence_{ij} + \beta_{5j} External\ influence_{ij} \\
&\quad + \beta_{6j} Government\ influence_{ij} + \beta_{7j}\ Formal\ coordination_{ij} \\
&\quad + \beta_{2j}\ Lateral\ coordination_{ij} + \beta_{3j} Informal\ coordination_{ij} \\
&\quad + \beta_{11j} File\ sharing\ technology_{ij} + \beta_{8j} Principal\ city_{ij} \\
&\quad + \beta_{9j} Socio - political\ barriers_{ij} + \beta_{10j} Technical\ barriers_{ij} \\
&\quad + \beta_{11j}\ Department\ size_{ij} + \beta_{13j}\ Tech\ capacity_{ij} \\
&\quad + \beta_{14j}\ Form\ of\ government_{ij} + \beta_{15j}\ Population\ (log)_{ij} \\
&\quad + \beta_{16j} Department\ type_{ij}
\end{aligned}$$

$\boldsymbol{Where}$ ($\boldsymbol{level\ 2}$):

$$\beta_{0j} = y_{00} + y_{01} Privacy\ laws_j + y_{02}\ Chief\ Data\ Officer_j + y_{03}\ Open\ data\ portals_j + U_{0j}$$

The first step to estimate a multilevel model is to calculate the intraclass correlation index (ICC). The ICC represents the proportion of variance explained by level-2 variables on the total variance of the model. In the case of a logit model, the variance for level-1 is fixed and corresponds to:

$$Var(\varepsilon|x) = \pi^2 / 3$$

Therefore, the ICC formula for a multi-level logit model is:

$$\sigma^2 v \ / \ (\sigma^2 v + (\pi^2 / 3))$$

To estimate the level-2 variance, I fit a null model (i.e., a model including only the intercept). The cluster number is 42. The average number of observations is 9.6 per cluster, ranging from a minimum of 1 observation to a maximum of 81 observations. The ICC for each model is reported in table 11.

The "non-governmental organizations" model has the highest ICC, while it the "other departments in the city" model has the lowest ICC. These values suggest that state-level variables account for more variation in inter-organizational data access than in intra-organizational data access. In both the "other public agencies" and the "non-governmental organizations" model, the ICC is high enough to justify a multilevel model. In the "other departments in the city" model, the ICC is relatively low for a multilevel model (Clarke & Wheaton, 2007).

Table 11 also reports the design effect and the actual sample size. The design effect indicates how many observations we would need to achieve the same N power, had the observations been independent of each other. For instance, in the "other departments in the city" model, a design effect equal to 1.5 means that we would need 1.5 times observations (Maas & Hox, 2005; Snijders & Bosker, 1993). The design effect ranges from 1.5 to 3.2. The actual sample size represents the size of the sample used for the estimation of the standard errors given the number of observations in the analysis and the design effect. The ICC, the design effect, and the actual sample size are correlated measures, such that a higher ICC corresponds to a higher design effect and a lower actual sample size.

Table 11.

Level-2 variance, ICC, and design effect of three multi-level data access models.

Model	Level 2 variance	ICC	Design effect	Actual sample size
Internal departments	0.205	0.059	1.504	267.28
Public agencies	0.594	0.153	2.316	173.59
Non-governmental orgs	1.140	0.257	3.213	125.12

First, I estimate a model with only level-1 variables; then, I included level-2 variables. The model has a random intercept and fixed slopes. Table 12 reports the results. Significant results ($p < 0.1$) are bold. I also performed a Wald test to compare an unrestricted model where all β coefficients can vary with a restricted model where level-2 variable coefficients are set equal to zero. Because weights are used in the analysis, a

Wald test is more appropriate than a Likelihood Ratio test. The Wald test suggests that the full model is an improvement compared to the null model ($p<0.001$).

Table 12.

Data access from other departments in the city, other public agencies, and non-governmental organization - Multilevel model estimator.

	Data access from other depts. in the city				**Data access from other public agencies**				**Data access from nongovernmental organizations**			
	Beta	**Odds ratio**	**SE**		**Beta**	**Odds ratio**	**SE**		**Beta**	**Odds ratio**	**SE**	
Institutions												
Privacy Laws	0.00	1.00	0.01		-0.01	0.99	0.01		-0.01	0.99	0.01	
Legal Mandate	**0.38**	**1.46**	**0.16**	*	**0.35**	**1.42**	**0.15**	*	**0.23**	**1.26**	**0.13**	+
Institutional Capacity	0.04	1.04	0.07		0.08	1.08	0.08		**0.15**	**1.16**	**0.09**	+
Open Data Portals	0.08	1.08	0.12		-0.19	0.83	0.17		-0.14	0.87	0.15	
Social environment												
City Influence	0.14	1.16	0.18		-0.03	0.97	0.13		-0.02	0.98	0.13	
Civil Society Influence	0.28	1.33	0.21		**0.54**	**1.72**	**0.20**	**	**0.48**	**1.61**	**0.22**	*
Government Influence	-0.29	0.75	0.20		**-0.41**	**0.67**	**0.15**	**	**-0.37**	**0.69**	**0.16**	*
Coordination mechanisms												
Formal Coordination	0.10	1.10	0.16		0.24	1.27	0.15		0.17	1.19	0.19	
Lateral Coordination	-0.13	0.88	0.16		0.04	1.04	0.19		0.23	1.26	0.17	
Informal Coordination	**0.48**	**1.62**	**0.16**	**	**0.40**	**1.49**	**0.18**	*	**0.29**	**1.33**	**0.17**	+
Control variables												
Principal City	**0.85**	**2.34**	**0.31**	**	**0.54**	**1.72**	**0.32**	+	**0.69**	**2.00**	**0.34**	*
Social Barriers	-0.40	0.67	0.24		-0.46	0.63	0.23	*	-0.31	0.74	0.19	
Technical Barriers	0.20	1.23	0.23		0.15	1.16	0.23		0.16	1.18	0.18	

File Sharing Technology	**0.41**	**1.51**	**0.25**	+	0.18	1.20	0.27		**0.80**	**2.22**	**0.23**	***
Dept Size	0.00	1.00	0.10		-0.09	0.92	0.12		-0.04	0.97	0.11	
Technical Capacity	**-0.36**	**0.70**	**0.21**	+	-0.19	0.82	0.18		**-0.35**	**0.70**	**0.16**	*
Form of Government	-0.19	0.83	0.27		-0.15	0.86	0.25		0.16	1.18	0.34	
Population	**-0.39**	**0.68**	**0.19**	*	-0.16	0.86	0.20		**-0.39**	**0.68**	**0.16**	*
Community Development	-0.62	0.54	0.43		**-0.80**	**0.45**	**0.44**	+	-0.22	0.80	0.44	
Finance	-0.68	0.50	0.55		-0.61	0.54	0.50		-0.19	0.83	0.43	
Parks and Recreation	-0.78	0.46	0.49		-0.65	0.52	0.49		**-0.61**	**0.55**	**0.37**	+
Police	0.57	1.77	0.58		**0.93**	**2.55**	**0.48**	*	**0.80**	**2.23**	**0.37**	*
Cut 1	-3.71		2.65		-1.75	0.17	2.36		**-3.55**	**0.03**	**1.84**	+
Cut 2	-2.18		2.59		0.08	1.09	2.30		-1.67	0.19	1.82	
State variation	0.41		0.40		0.72	2.05	0.78		1.04	2.82	0.58	
Observations		402				402				402		
Clusters		42				42				42		
Averaged cluster		9.6				9.6				9.6		
Log likelihood		-836.2109				-895.2611				-925.7676		
AIC		1722.422				1840.522				1901.535		
BIC		1822.333				1940.433				2001.447		

Significance levels: + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ - Reference category: Mayor’s Office

Seemingly Unrelated Regression Model

Up to now, each model was estimated separately. However, it is reasonable to assume that the three equations are correlated, as the same respondent has reported data access across the whole portfolio of relationships. Moreover, we might expect that departments which are better at accessing data from other city departments might also be better at accessing data from other organizations. Empirically, the three dependent variables are highly correlated (>0.5) suggesting that there might be common factors explaining their variation (correlations are reported in table 13). If that is the case, the error terms of the regressions are correlated, and a seemingly unrelated regressions (SUR) estimator is preferable.

The SUR estimator was developed by Zellner (1963) and accounts for the likelihood that the error terms of two or more regressions are not independent. If the error terms are uncorrelated, then the SUR estimator is not different from other estimators. But, when the error terms are correlated, we gain a more efficient estimator by jointly estimating the equations and taking into account the full covariance structure (Greene, 2000; Zellner; 1963)[11].

[11] The SUR estimator provides efficiency gains if the matrix of the observations differs across equations (e.g., at least one variable is different across equations). In the presented models, the variable Legal Mandate is measured differently across the three equations such that it is possible to apply a SUR estimator.

Table 13.

Correlation across dependent variables

Data access from	Other depts. in the city	Public agencies	Non-governmental organizations
Other depts. in the city	1.00		
Other public agencies	0.76	1.00	
Non-governmental organizations	0.62	0.72	1.00

The SUR model was estimated using the "cmp" command in Stata (Roodman, 2011). "cmp" is a flexible statistical framework, which relies on a maximum likelihood (ML) estimator to estimate systems of two or more equations across a wide range of variable types. Given the nature of the DVs, I utilize an ordered model with a probit link[12]. Design weights and clustered robust standard errors are applied to each regression as described in previous sections. A SUR multilevel model can also be estimated in "cmp" but unfortunately, the model does not converge. Non-convergence is common in seemingly unrelated multi-level models given the number of parameters that require estimation (Roodman, 2011). To try to reach convergence, I estimate the model first, assuming that all error terms are correlated and then, assuming that only error terms at level-1 are correlated; this latter option generally facilitates the estimation of the parameters. In both cases, the model did not converge.

Results from the SUR model are shown in table 16. Odd ratios are not reported because the estimation is based on a probit link. Probit coefficients represent a change in

[12] The logit link is not available under the cmp package.

the z-score of the underlying variable (Long, 1997). Significant coefficients ($p < 0.1$) are bold. Correlations across the equations are reported in table 14. The rhos[13] – the correlations of the residuals between each pair of regressions - are significant for all pair of regressions, confirming that error terms are correlated. A series of Walt tests comparing the full model with a constrained model where the correlation across error terms is set, first equal to zero for each pair of regression, and then equal to zero for all pairs of equations, confirm the appropriateness of a SUR system ($p<0.001$). Results of the Wald tests for each applied constraint are reported in table 15.

Table 14.

Correlations across regression in the SUR model presented in table 16.

	Beta	SE	P-value		Confidence Interval	
/atanhrho_12	1.36	0.11	0	***	1.15	1.57
/atanhrho_13	0.94	0.10	0	***	0.74	1.14
/atanhrho_23	1.23	0.11	0	***	1.00	1.45
rho_12	0.88				0.03	0.92
rho_13	0.74				0.05	0.82
rho_23	0.84				0.03	0.90

Table 15.

Applied constraints to the SUR system in table 16 and Wald test results.

Constraints	**Chi2**	**DF**	**Prob > chi2**
Rho_12 = 0	158.47	1	0.000
Rho_13 = 0	82.41	1	0.000
Rho_23 = 0	118.58	1	0.000
Rho_12, Rho_13 and Rho_23 = 0	283.94	3	0.000

[13] The cmp estimates the Fisher's z transformed rho

Table 16.

Data access from other departments in the city, other public agencies, and non-governmental organization – Seemingly Unrelated Regression (SUR) model

	Data access from other depts. in the city			**Data access from other public agencies**			**Data access from nongovernmental organizations**		
	Beta	**SE**		**Beta**	**SE**		**Beta**	**SE**	
Institutions									
Privacy Laws	-0.00	0.00		-0.00	0.00		-0.00	0.01	
Legal Mandate	**0.17**	**0.06**	**	**0.13**	**0.04**	***	**0.16**	**0.06**	**
Open government	0.02	0.05		**-0.09**	**0.05**	+	-0.06	0.06	
Institutional capacity	0.00	0.02		0.02	0.02		0.04	0.02	
Social environment									
City Influence	0.03	0.10		-0.04	0.07		0.01	0.07	
Civil Society Influence	**0.28**	**0.13**	*	**0.41**	**0.09**	***	**0.31**	**0.10**	**
Government Influence	**-0.18**	**0.10**	+	**-0.25**	**0.09**	**	**-0.22**	**0.08**	**
Coordination mechanisms									
Formal Coordination	0.08	0.08		**0.16**	**0.09**	+	0.11	0.10	
Lateral Coordination	-0.06	0.09		0.07	0.10		**0.17**	**0.09**	+
Informal Coordination	**0.26**	**0.09**	**	**0.20**	**0.09**	*	0.13	0.10	
Control variables									
Principal City	**0.47**	**0.15**	**	**0.29**	**0.17**	+	**0.41**	**0.20**	*
Social Barriers	**-0.24**	**0.13**	+	**-0.27**	**0.12**	*	**-0.19**	**0.09**	*
Technical Barriers	0.10	0.13		0.03	0.12		0.05	0.08	
Dept Size	-0.02	0.05		-0.06	0.06		-0.03	0.06	

Technical Capacity	**-0.20**	**0.11**	+	-0.14	0.09		**-0.23**	**0.08**	**
File Sharing Technology	0.20	0.14		0.07	0.16		**0.37**	**0.15**	**
Form of Government	-0.14	0.15		-0.01	0.13		0.11	0.17	
Population	**-0.27**	**0.08**	***	-0.13	0.09		**-0.23**	**0.09**	*
Community Development	-0.33	0.22		-0.48	0.20	*	-0.16	0.23	
Finance	-0.29	0.28		-0.27	0.25		-0.11	0.22	
Parks and Recreation	**-0.40**	**0.24**	+	-0.34	0.26		-0.29	0.20	
Police	0.37	0.29		**0.58**	**0.24**	*	**0.48**	**0.21**	*
Cut 1	**-2.84**	**1.27**	*	-1.56	1.16		**-2.15**	**0.97**	*
Cut 2	-2.03	1.25		-0.61	1.14		-1.14	0.98	

Significance levels: + p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001, Reference categories: Mayor’s Office

Summary

I estimated three different models to assess the most appropriate fit to the data and check the robustness of the results. I used a probit model with clustered SEs as the baseline model. I then estimated two other models: a multilevel model and a system of seemingly unrelated regressions model. Here I briefly summarize the fit of each model before proceeding to discuss the results in the light of the hypotheses outlined in chapter 4.

The clustered SEs model shows that an ordered logit estimator is appropriate as the parallel line assumption is met. Only four variables were found problematic. A partial non-proportional odds model suggests that there are only small differences across categories to be considered in the interpretation of the findings.

The multilevel model considers the two-level structure of the data and facilitates inference on level-2 variables. It considers that the variance of the dependent variable is explained by both department (level-1) and state (level 2) factors. The ICC is low in the "other departments in the city" model (ICC < 0.1), but it is moderated in the "other public agencies" (0.1 > ICC > 0.2) and "non-governmental organizations" model (0.2 > ICC > 0.3).

However, the multilevel model presents some limitations because of design issues related to group size. The estimated models include 42 clusters of 1 to 81 observations, for an average cluster size of 9.6. Sparseness - the presence of clusters with only one observation - might cause bias in model estimation (Bell, Ferron, & Kromrey, 2008). In the data, 19% of clusters are singletons, i.e., they contain only one observation (8 clusters out of 42).

Sparseness is a common problem in social sciences (Bell, Ferron, & Kromrey, 2008). Singletons might cause problems in a multilevel structure because the individual and the within-group variability are equal when a cluster is composed of only one individual. Eliminating singletons might create even more problems if singletons are systematically different from other clusters. Several studies have tried to assess the impact of singletons and small cluster size on multilevel estimation using simulations. Most of them find that sparseness does not affect fixed-effect coefficients (Bell, Ferron, & Kromrey, 2008; Clarke & Wheaton, 2007; Maas & Hox, 2005) but it does affect random-effect estimates. These results indicate that the group-level variance - which is of interest in this research - can be overestimated. Over-estimation of the group-level variance increases when the model has a low ICC (i.e., ICC = 0.1) and a small number of groups (<100) [14] (Clarke & Wheaton, 2007). These conditions might cause Type I

[14]Biases in multilevel model results is one reason why this research does not estimate random slope parameters. The bias created by small group size and sparseness is accentuated when random slope parameters are estimated as compared to models that include only random intercept parameters (Clarke & Wheaton, 2007).

statistical error and lead to infer that there are group-level differences when there are none[15,16].

In this research, singletons are a reasonable percentage of the clusters, the model has relatively low ICC, and the number of clusters is low. Because of these issues, I suggest that a multilevel model should be utilized to validate the baseline ordered logit model, but we need to be cautious in drawing conclusions about the study. When compared to the baseline model, the standard errors of the multilevel model are more substantial, but the size and sign of coefficients are comparable.

Finally, the SUR model accounts for the correlations across the three equations. Empirically, the SUR system has the smallest standard errors, which can be explained by the fact that it includes more information in the estimation than other models. However, results do not substantially differ from the clustered SE model and the multilevel model. Given that the SUR model shows strong evidence that the error terms are correlated and results do not significantly differ from previous models, the next section discuss the hypotheses based on results from the SUR model. Appendix G and H discuss robustness check that have been conducted on the final model. Appendix G re-estimates the three model only considering cities with population below 100,000. Appendix H shows the

[15] Results were generated by simulating models with high sparseness (60% of singletons), medium sparseness (10% of singletons) and low sparseness (2% of singletons) (Clarke & Wheaton). The level of sparseness in this model (=0.2) suggests that bias will not be as large as in high sparseness model, but it won't neither be reduced as low as in the models with medium or low sparseness. Comparing the study's conditions with simulations holding similar characteristics, there is evidence of upward bias in random effects.

[16] Clarke and Wheaton (2007) at least ten observations per cluster and at least 100 clusters for the intercept variance to approach the true value.

estimation of the model when utilizing multiple imputation technique to impute missing values.

Hypotheses Testing

In this section, I discuss the results against the hypotheses presented in Chapter 4. I am mostly interested in the significance and direction of the correlations between the independent variables and the three dependent variables – data access from other departments in the city, other public agencies, and non-governmental organizations. The significance of the coefficients is defined at the 0.05 level, but coefficients with a p-value lower than 0.1 are also discussed. While 0.05 is the most commonly accepted significance level, it is not unusual to indicate results that meet the 0.1 threefold. These results still indicate a high probability of correlation and provide insights into managerial practices. I will also discuss the magnitude of the coefficients across variables or models; particularly, to facilitate the interpretation of the results from a probit model, I report the discrete change, that is "the change in the predicted probability for a change in x_k from a start value x_s to the end value x_e" (Long, 1997, p. 136). Predicted probabilities will focus on the "high data access" category and will describe the probability that department heads report high data access, holding all other variables constant at their means. All comparisons are based on a Wald test between standardized coefficients. Results are based on the Seemingly Unrelated Regression model (table 16). As mentioned above, results are consistent with the ordered logit model with clustered robust standard errors and the multilevel model.

Institutions. Hypothesis 1a states that privacy laws will be positively correlated with data access. The analysis finds no evidence that privacy laws affect data access from other departments in the city, other public agencies, or non-governmental organizations; the Privacy Laws coefficients are statically non-significant (p-value > 0.05). This finding is in contrast with previous research arguing that privacy laws either positively affect data access – because of reduced privacy and data misuse concerns - or negatively affect data access – because of higher requirements and constraints and fear of noncompliance.

It might be that privacy laws only affect data access, when data contain sensitive information, such as personal identifiers, social security numbers, addresses, medical records, and so on (Pew Charitable Trusts, 2018). To rule out this explanation, I introduced an interaction term between Privacy Law and Sensitive Data[17] to test whether Privacy Laws have any significant effect on data access when city departments exchange sensitive data. Table 17 shows the coefficients for the direct effect and the interaction terms across all three models (standard errors are reported into brackets). The full model is presented in Appendix F.

Table 17.

Results of the interaction between Privacy Laws and Sensitive Data.

	Other depts. in the city	**Other public agencies**	**Non-governmental orgs.**
Privacy Laws	.018 (.015)	.005 (0.011)	.012 (0.009)
Sensitive Data	.087 (0.278)	-.204 (0.260)	.174 (0.238)
Privacy Laws * Sensitive Data	-.005 (0.004)	.000 (0.004)	-.006 (0.003) +

[17] Sensitive Data is a continuous variable ranging from 1 (=strongly disagree) to 5 (=strongly agree). The survey item asked: "Most activities in my organization requires access to sensitive data that contains personally identifiable information" (mean = 2.99, s.d. = 1.17).

Significance levels: + $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Results show no significant correlation in the "other departments in the same city" and "other public agencies" models, but the coefficient is slightly significant and negatively correlated with data access (p-value = 0.06) in the "non-governmental organizations" model. This result suggests privacy laws might decrease city departments' capacity to access sensitive data from non-governmental organizations.

Hypothesis 1b states that legal mandates to share data will be positively correlated with data access. Across all models, Legal Mandate has a positive and significant coefficient, showing support for the hypothesis (p-values < 0.05). In the "other departments in the same city" model, moving from “no legal mandates” to “all requests are subject to legal mandates to share data” changes the predicted probability of reporting high data access by 0.19, holding all other variables constant at their means. In the "other public agencies" model, moving from “no legal mandates” to “all requests are subject to legal mandates to share data” changes the predicted probability of reporting high data access by 0.16, holding all other variables constant at their means. Finally, in the "nongovernmental organizations" model, moving from “no legal mandates” to “all requests are subject to legal mandates to share data” changes the predicted probability of reporting high data access by 0.18, holding all other variables constant at their means.

Results also show that Legal Mandate has a larger effect in the "other departments in the same city" and "other public agencies" models (Wald test, $p < 0.05$). The effect is lower in "non-governmental organizations" model.

Hypothesis 1c argues that state governments’ institutional capacity will be positively correlated with data access. Institutional capacity is not significant in any model. Therefore, there is no support for hypothesis 1c. Hypothesis 1d states that the quality of open data portals will be positively correlated with data access. The variable Open Data Portals is not significantly correlated with data access in any model (p-value > 0.1), thereby providing no evidence in support of hypothesis 1d. Taking together findings from hypothesis 1c and 1d, I can conclude that state government institutions - as measured in this study - are not significant predictors of data access capacity among city departments. Legal mandate is the only significant predictor among institutional variables.

Social environment. Hypothesis 2a suggests that the influence of external organizations will be negatively correlated with data access. When in a position of power, organizations might refuse to provide data to maintain their influence over city departments' decision-making processes.

I find similar results across all model models. City Influence is not significantly correlated with data access (p-value > 0.1) while Civil Society Influence has a positive effect on data access ($p < 0.01$). In fact, moving from “no influence” to “strong influence” of civil society actors increases the predicted probability of reporting high data access from nongovernmental organizations by 0.46 in the “nongovernmental organization model” and by 0.58 in the “other public agencies” model, holding all other variables constant at their means. This latter result contradicts hypothesis 2a. By contrast, Government Influence has a negative effect on data access, which supports hypothesis 2a (0.05 > p-value > 0.01). For instance, in the "other public agencies" model, moving from “no influence” to “strong influence” of government actors decreases the predicted probability of reporting high data access by 0.32, holding all other variables constant at their means. In the "nongovernmental organizations" model, the predicted probability decreases by 0.30. Because of mixed evidence, it is difficult to confirm hypothesis 2a.

Coordination mechanisms. Hypothesis 3a suggests that Formal Coordination will be positively correlated with data access from other public agencies. I find weak support for hypothesis 3a. The Formal Coordination variable is only slightly significant and positively correlated with data access from public agencies in the SUR model (p-value < 0.1). The predicted probability of reporting high data access increases by 0.24, holding all other variables constant at their means when we move from the minimum value of formal coordination to its maximum.

Hypothesis 3b suggests that Lateral Coordination is positively correlated with data access from non-governmental organizations. I find a positive and significant correlation between Lateral coordination and data access from non-governmental organizations (p-value = 0.05) in the SUR model. The predicted probability of reporting high data access increases by 0.26, holding all other variables constant at their means when we move from the minimum value of lateral coordination to its maximum. However, it is worth noted that this result is not consistent in the multilevel model (p>0.05).

Hypotheses 3c and 3d suggest that Informal Coordination is positively correlated with data access from other departments in the city and non-governmental organizations, respectively. Results show support for hypothesis 3c but provide limited evidence in support of hypothesis 3d. All models show a significant and positive correlation between informal coordination and data access from city departments ($p<0.05$). The predicted probability of reporting high data access from other departments in the city increases by 0.4 when comparing department heads who never utilize informal coordination with department heads who always utilize informal coordination, holding all other variables constant at their means. However, the correlation between informal coordination and data access from non-governmental organizations is non-significant at the 0.05 level. Unexpectedly, I find a positive and significant relationship between lateral coordination and data access from other public agencies. The predicted probability of reporting high data access from other public agencies increases by 0.31 when comparting department heads who never utilize informal coordination to department heads who always utilize informal coordination, holding all other variables constant at their means.

Control variables. Among control variables, I find no evidence that Technological Barriers decrease data access; the coefficient is statistically not significant (p-value > 0.05). By contrast, Socio-political Barriers are negatively and significantly correlated with data access from other public agencies (p-value < 0.05), non-governmental organizations ($p<0.1$), and other departments in the city ($p <0.1$).

Technical Capacity has a negative and significant effect on data access from non-governmental organizations (p-value < 0.05). Technical Capacity is also slightly significantly correlated with data access from other departments in the city (p-value < 0.1). I also find that information sharing technologies are positively correlated with data access. All models show a positive relationship, but the variable File Sharing Technology is significantly correlated with data access only in the "non-governmental organizations" model (p-value < 0.01).

Results also show that departments in principal cities are more likely to access data from other departments in the city and non-governmental organizations. A slight significant correlation is also found in the data access from public agencies model (p-value = 0.1).

The analysis included five department types: Mayor’s office, community development, finance, parks and recreation, and police. The Mayor’s office is the reference category in all models. Results show variation across department types. Police department heads report higher data access from public agencies and non-governmental organizations than Mayor’s office heads. Community development department heads are less likely to access data from other public agencies as compared to Mayor’s office heads.

Finally, findings show a significant effect of Population Size, where department heads working for smaller cities report lower data access from other departments in the city and non-governmental organizations (p-values < 0.05). This finding is further explored in Appendix E where I present a separate model for small cities.

CHAPTER 7

DISCUSSION AND CONCLUSIONS

Following the analysis conducted in chapter 6, this chapter discusses the main findings and conclusions from this research. The chapter is organized in four sections. The first section discusses the results of the statistical analysis. The section also highlights the main theoretical contributions of the study. The second section discusses data and methodology limitations, including common method bias, measurement issues, and survey data. The third section summarizes contributions to managerial practices and key takeaway points for public managers who wish to support data practices in general, and particularly data access. Finally, the last section proposes opportunities for future research in public management and data sharing scholarship.

Discussion of Results

This research focuses on data access for several reasons. First, despite the development of new technology for collecting, analyzing, and storing data, public managers greatly rely on data provided by nonprofit and for-profit organizations and other public agencies to feed decision-making processes and support the design and development of public policies (Allard et al., 2018; Pew Charitable Trusts, 2018). Public managers expect that greater availability and diversity of data will help them to improve public outcomes, including efficiency, effectiveness, and equity (Allard et al., 2018; Jennings & Hall, 2012). Moreover, data are raw inputs for the creation of novel information and knowledge, which can promote and support innovation in the public sector. Therefore, understanding the determinants of data access is vital to improve government decisions and the quality of public services for the benefit of citizens. Second, given the growing complexity of the policy environment, public organizations are required to engage in some form of data, information, and knowledge sharing with external and internal stakeholders. Sharing data and information is needed to collaborate with other organizations, coordinate common activities, and take into account different perspectives when addressing complex policy problems. Yet, recent research suggests that public managers still face severe barriers to access and exchange data and information (Allard et al., 2018; Feeney et al., 2016). Additional research is needed to increase the ability of public managers to obtain data and information and reduce time, human, and financial resources that are needed to bend data access barriers.

This research develops an Integrative Framework for Data Access (IFDA) and uses it to investigate data access across departments in small- and medium-sized cities in the United States. The IFDA combines institutional theory, resource dependence theory, and collaboration studies to highlight factors that influence data access across various levels of analysis. The IFDA assumes that public managers initiate efforts to access data from other organizations through different coordination mechanisms and that the institutional and social context of city governments shape opportunities and constraints for data access. Moreover, the IFDA considers institutional and social differences that characterize the portfolio of stakeholders with whom city departments exchange data - other departments in the city, other public agencies, and nongovernmental organizations.

The IFDA brings three major contributions to public management research and data sharing literature. First, previous studies have suggested the appropriateness of a multilevel model to investigate data and information sharing (Fountain, 2007; Yang & Maxwell, 2011) but public management scholarship lacks a fully developed theoretical model followed by an empirical investigation. The IFDA acknowledges and empirically tests how factors at multiple levels – institutions, social relationships, and coordination – affect data access; these factors are analyzed separately in previous research.

Second, it expands previous studies that mostly focus on collaborative initiatives for data sharing. These studies look at situations where public organizations collaborate with their stakeholders to set up rules, structures, and incentives to share data. However, in their daily activities, public managers are likely to interact with stakeholders that are autonomous and have no or few incentives to share data. This research looks at data sharing more broadly and emphasizes the daily coordination mechanisms that public managers utilize to share data across their portfolio of relationships. By doing so, it examines how formal, lateral, and informal coordination mechanisms coexist to facilitate data access and the conditions under which coordination is effective.

Particularly, the IFDA focuses on how coordination mechanisms vary across stakeholder types and highlights social and institutional differences across the portfolio of relationships of a city department. Based on previous literature (Daley, 2009; Eglene, Dawes, & Schneider, 2007), I argue that there are differences between intra- and inter-organizational relationships and same and cross-sector ones. If scholars and public managers wish to increase data access across public agencies, then it fundamental to understand such differences and investigate data access across the full relationship portfolio to inform research and practice.

I test the IFDA using data collected by the Center for Science, Technology, and Environmental Policy Studies at Arizona State University. The survey collects data about 2,500 city departments in 500 US cities with populations from 25,000 to 250,000. Descriptive statistics show that city departments have relatively good access to data, but there are significant differences across the portfolio of relationships. In line with previous research (Dawes, Cresswell, & Pardo, 2009; Roberts, 2011), city departments report lower data access from non-governmental organizations and greater access from other departments in the city. Technical and organizational barriers are higher when exchanging data across sectors, and sectoral differences might exacerbate socio-political and cultural conflicts associated with data misuse or the negotiation of conditions for using and accessing data (Susha, Janssen, & Verhulst, 2017).

Results from the analysis show only partial support for most hypotheses of the IFDA, as summarized in table 18. Macro-level factors, such as legal mandates and stakeholder influence, have a similar effect across the portfolio of stakeholders, while coordination mechanisms substantially differ based on the stakeholder type. I discuss the results in more detail in the next paragraphs.

Table 18.

Summary of research results.

Hypotheses	***Results***
Institutions	
H1a. State privacy laws will be positively correlated with data access.	Not supported
H1b. A legal mandate to share data across organizations will be positively correlated with data access.	Supported
H1c. State government's institutional capacity will be positively correlated with data access.	Not supported

H1d. Quality of state government open data portals will be positively correlated with data access.	Not supported
Social environment	
H2a. Influence of external organizations will be negatively correlated with data access.	Partially supported • Influence from city actors – Not supported • Influence from non-governmental organization – Significant but positively correlated • Influence from other governmental agencies - Supported
Coordination mechanisms	
H3a. Formal coordination mechanisms will be positively correlated with data access from other public agencies.	Not supported
H3b. Lateral coordination will be positively correlated with data access from nongovernmental organizations.	Not supported
H3c. Informal coordination mechanisms will be positively correlated with data access from other departments in the city.	Supported
H3d. Informal coordination mechanisms will be positively correlated with data access from nongovernmental organizations.	Not supported

Institutions. The IFDA argues that institutions are a fundamental part of the environment in which data exchange takes place, and they might limit or facilitate data access. Institutions are humanly devised constraints and incentives that exert normative and coercive pressures on organizations and individuals to act accordingly to socially accepted behaviors and rules. Results show that only coercive pressures in the form of legal mandates to share data matter for explaining data access. I find weak to no evidence that other coercive pressures, such as privacy laws, and normative pressures, such as state institutional capacity and the quality of open data portals, are related to data access. I advance some explanations for these results.

Previous research suggests that privacy laws either increase (Fountain, 2007; Yang & Maxwell, 2011) or decrease data sharing (6, Bellamy, Raab, Warren, & Heeney, 2007). Findings from this study suggest an alternative hypothesis: while privacy laws establish an overarching framework for sharing data across organizations, they do not directly affect the decision to share or not to share data with public organizations. Since privacy laws are relatively weak in the United States (Baumer, Earp, & Poindexter, 2004), it might be that managers are little concerned with noncompliance and, therefore, privacy laws do not prevent organizations from providing data to public agencies. Future studies might compare the effect of privacy laws between countries with weak requirements, such as the US, and countries with strong requirements, such as member states of the European Union. It might also be that managers are familiar with the few existing privacy laws and they know how to comply with them, so as privacy laws do not constitute a barrier for sharing data. Quasi-experimental studies could examine whether data sharing is restricted when states adopt a new privacy regulation and thereby discerning the effect of learning on institutional barriers.

Results show that legal mandates to share data have a positive effect on data access. This finding confirms evidence from previous case studies and interviews with public managers, which found that legal mandates greatly facilitate access to data across public agencies (6 et al., 2007; Allard et al., 2018; Jennings & Hall, 2012). This study further suggests that legal mandates have a greater effect on data exchange among departments in the same city and public agencies. I found a lower effect when examining data access from nongovernmental organizations. It might be that state governments are less involved in cross-sectoral data exchange and are less likely to force nongovernmental organizations to share data with city governments. By contrast, state governments are increasingly regulating data sharing across public organizations in an attempt to break down departmental silos (Pew Charitable Trusts, 2018).

The interpretation of the results on legal mandates should consider that the study focuses on "timely" data access - i.e., whether public managers can access data without the need to follow up with the other organization. It might be that legal mandates to share data have a lower impact on the responsiveness of nongovernmental organizations. Because they less frequently exchange data with city departments (Feeney et al., 2016), non-governmental organizations might need time to put together data requested by city governments; learn about institutional constraints; address organizational policies and requirements; and obtain authorization to share data. Thereby, public managers might be able to access data only after submitting follow-up requests. Future research should investigate how legal mandates impact on the overall capacity to access data from nongovernmental organizations and the quality of data that the organizations provide (see: measurement limitations).

Regarding normative pressures, I do not find a significant correlation neither in the case of institutional capacity nor quality of the open data portals. These results contrast with previous research on e-government and information technology adoption (Grimmelikhuijsen & Feeney, 2016; McNeal, Tolbert, Mossberger, & Dotterweich, 2003b), collaboration (Smith, 2009), and information sharing (Dawes et al., 2009). A possible explanation regards the level of analysis at which institutions were measured.

Some studies on institutional theory found that normative and coercive pressures stem at different levels (Jennings & Zandbergen, 1995). While coercive pressures stem from high-level institutions - such as state governments -, peer organizations and organizations in the same field are the primary sources of normative pressures.

According to this perspective, it might be that state-level normative pressures have no effect on data exchange at the local level. Internal and external stakeholders refer to the institutional capacity and culture of the city government when looking for social cues and cultural norms to understand acceptable practices and behaviors. For instance, stakeholders might look at whether the city government provides an open data portal or whether it has a data sharing policy. Policy and innovation diffusion theories (Berry & Berry, 2014) corroborate this perspective, suggesting that organizations tend to adopt behaviors that are similar to those of their peers. Future research should examine the effect of social cues at the local level, including symbolic and cultural signals that city governments might diffuse in the environment through websites, social media, and open data portals.

Social environment. This study recognizes that city department heads are not isolated agents; requests for data occur within a set of previously established relationships (Azad & Wiggins, 1995). The IFDA draws from resource dependence theory (Pfeffer & Salancik, 1978) and focuses on power dynamics, which are a fundamental component of data exchange. Issues of dependence, political power, and influence across organizations are likely to shape willingness to share data (Gil-Garcia, Pardo, & Nam, 2016; Meijer, 2018). I look at influence, defined as the ability of external organizations to pressure city departments to address their own and their constituency's interests. Because of their size, small cities are greatly subject to influence from local constituencies (Hamin, Gurran, & Emlinger, 2014).

Resource dependence theory suggests that data are a critical resource for organizations. The decision to hide or disclose data and information is a strategic and political one, and organizations might refrain from sharing them in order to maintain power in the environment, ensure a competitive advantage, and avoid damages to their performance. I find mixed evidence in support of this hypothesis. Overall, the analysis shows that influence dynamics broadly affect data access across the full portfolio of relationships, but there are relevant differences across stakeholder types. While state and federal government influence reduces data access, civil society influence has a positive effect, and city actor influence – the Mayor, the city council and other city departments – does not affect data access. I discuss three main implications for theory and practice.

First, city governments that are more intermingled into sectoral power dynamics experience negative repercussions in engaging with internal and external stakeholders to access data. The relationship between city departments and the state and the federal government is likely characterized by hierarchy, conflicts over shared authority, legal requirements, bureaucratic culture, and enforced accountability across agencies (Fountain, 2007; Mullin & Daley, 2010). When government agencies are highly influential, city departments might experience lower autonomy and discretion to negotiate conditions for data sharing, which in turn negatively affect data access. Bureaucratic rules and hierarchical relationships might also constrain public managers' action and reinforce jurisdictional divisions and accountability requirements across departments and agencies, which decrease willingness to share data. This finding suggests that the bureaucratic form of government is still persistent in the public sector and reinforce issues of accountability and rule enforcement that bind opportunities for sharing and collaborating across organizations (Dawes, 1996; Fountain, 2007; Roberts, 2011).

Second, I suggest that there are different explanations as of why civil society influence increases data access from public sector organizations - both other departments in the city and other public agencies - and nongovernmental organizations, respectively.

Open data literature suggests that pressures from the civil society positively affect government willingness to provide data (Grimmelikhuijsen & Feeney, 2016). Government agencies are responsive to civil society pressures because they need to maintain political legitimacy and draw support for policy adoption and implementation (Rainey, 2009). Cities that experience greater influence from civil society organizations might need to collaborate more closely with other government actors in order to respond to demands from citizens, advocacy groups, professional associations, and media. By pressuring the government, civil society creates an environment that is more conducive of collaboration and data and information sharing across government agencies, which increases the likelihood that department heads will obtain data.

Civil society influence also increases data access from non-governmental organizations. Non-governmental organizations might be more willing to provide data when they perceive that external stakeholders can influence policymaking. Providing data might be a way to maintain and reinforce influence and shape policy decisions. Additionally, when civil society is highly influential, department heads cannot ignore their inputs and need to acquire more information to develop policies and services that meet their interests. Thereby, they might be more willing to engage in negotiations and accommodate nongovernmental organizations' requirements regarding data use and exchange. Given these findings, future research should further investigate the quality of data that city departments are timely able to obtain from non-governmental organizations. Non-governmental organizations might provide only data that support their policy preferences and benefit their constituencies. If that is the case, public managers should pay attention to how information available might affect decision making and its impact on outcomes for citizens, especially less represented communities.

The flows of financial resources might also explain the relationship between city departments and non-governmental organizations. City departments are likely to provide funding to nongovernmental organizations or develop contracting relationships for the provision of public services – e.g., public-private partnerships. Funding and contracts might come with requirements of accountability and transparency – e.g., accountability clauses - that force non-governmental organizations to regularly provide data to city governments (Malatesta & Smith, 2012; Reynaers & Grimmelikhuijsen, 2015).

Finally, city actor influence is not a significant predictor of data access. It is reasonable to assume that internal power dynamics have little effect on the outcome of data exchange with other public agencies and nongovernmental organizations. Department heads are the ones in charge of dealing with external stakeholders, while the

Mayor and the city council are not involved in the day-to-day interactions with external organizations. By contrast, it is surprising that internal influence does not affect the exchange of data across departments in the same city. Previous studies have found that internal power dynamics affect knowledge exchange across organizational units (Willem & Buelens, 2007). The small size of the cities included in the sample might explain this result. In small- and medium-sized, the small number of council members and departments might encourage internal communication and coordination, which moderate the effect of influence dynamics.

Coordination mechanisms. Results regarding coordination mechanisms (formal, lateral, and informal) provide only partial support for the hypotheses, but show some interesting findings that shed light on how city department heads can access data from internal and external stakeholders. While formal and lateral coordination are slightly correlated with data access, informal coordination significantly explains data access across the portfolio of relationships.

Previous research on collaboration and information sharing has suggested that routines facilitate the exchange of resources across organizations (Bryson, Crosby, & Stone, 2006; Dawes et al., 2009; Simo & Bies, 2007). However, I find that formal coordination has a weak and positive relationship only with data access from other public agencies. Small- and medium-sized cities might not be able to adequately design formal coordination mechanisms (e.g., routines and data sharing agreements) and they might encounter challenges in enforcing common rules. Previous studies have shown that organizations have strong preferences concerning data sharing routines and they might refrain from sharing data if routines do align with their organizational goals and values (Bekkers, 2007, 2009; Yang, Pardo, & Wu, 2014).

This finding opens several research questions for future studies, which should investigate the types of formal coordination mechanisms used by public managers to receive data from other organizations; how public managers design formal coordination and how other organizations participate in the process; and which challenges and barriers most likely affect the success of formal coordination. It is possible that different mechanisms for formal coordination will lead to different data access outcomes.

I find that lateral coordination has a positive, although weak, effect on access to data from nongovernmental organizations. This result provides weak evidence for hypothesis 3b but should be taken with caution. The interpretation of this result should consider that the study has no information on how nongovernmental organizations regulate the provision of data to other organizations. It might be that these organizations place constraints on sharing data with the city or grant little discretion to their employees to share data. Allard and colleagues (2107) report that public agencies have lengthy and tedious processes to share data, which fundamentally delay data access. Scholars should further investigate organization-level policies regarding data sharing and access to understand the effect of lateral coordination.

Previous studies have also shown that written formal requests for information are often declined or dismissed by employees unless coercive mechanisms, such as Freedom of Information regulation, support them (Worthy, John, & Vannoni, 2017). It is possible that similar motivations apply to this study, where employees dismiss written formal requests for data because they lack coercion. Future studies could explore the interaction between legal mandates and lateral coordination[18].

Finally, results show that informal coordination facilitates access to data from other departments in the same city and other public agencies. This finding corroborates previous research noting the importance of interpersonal networks for sharing data, information, and knowledge (6, Bellamy, Raab, Warren, & Heeney, 2007; Chen & Lee, 2018; Gil-Garcia, Pardo, & Burke, 2010; Kim & Lee, 2006; Nahapiet & Ghoshal, 1998; Willem & Buelens, 2007). However, I do not find that interpersonal networks matter when public managers seek access to data from non-governmental organizations. I discuss these results.

Same-sector relationships are characterized by similar norms and values that facilitate inter-personal collaboration and resource sharing (Fountain, 2007; Nahapiet & Ghoshal, 1998; Tsai, 2002). Public managers might rely on informal relationships to negotiate data access and collaboratively draw from professional training and norms to fill-in spaces left by organizational norms and rules (6 et al., 2007). Personal relationships might provide opportunities to work around legal and bureaucratic constraints as well as unclear regulations, thus facilitating access to data in the public sector. Instead, in cross-sector relationships, diversity of values, norms, and practices might be obstacles to informal coordination mechanisms (Bryson et al., 2006; Daley, 2009; Guo & Acar, 2005), such that public managers who more often utilize informal coordination do not

[18] Preliminary analysis that was conducted by the author with the available data shows no effect of an interaction term between lateral coordination and legal mandates. However, data are collected by macro categories; future research should consider better measurement by using network data.

report significantly greater access to data than the ones who rarely rely on informal coordination.

Additionally, public managers might be more likely to have ties with managers and employees working within their organization and in other public agencies. Public managers are more likely to collaborate and frequently interact with other actors in the public sector and therefore develop more extensive networks. Network composition might explain why interpersonal networks matter for same-sector relationships, but are not significant when public managers request data from nongovernmental organizations. This alternative explanation cannot be tested with available data; further research is needed to understand the breadth and composition of public managers' networks across stakeholder type.

Overall, results regarding coordination mechanisms suggest that data access is mostly a function of the initiative of city department heads rather than formal routines. Bellamy and colleagues (2008) reached a similar conclusion and argued that "achieving the enhanced volume of information-sharing demanded by government policy therefore relies, in practice, on the persistence of individual-goal seeking behaviors enabled by individualist forms and on the coping mechanisms" (p. 756).

Control variables. It is worth highlighting some results from the analysis of the control variables that hold implications for future research and practice. I focus the attention on five variables: technological barriers and file sharing technologies; socio-political barriers; principal cities; and department type.

First, technological barriers are not correlated with data access in any model. This result contradicts previous research suggesting that technical barriers fundamentally affect data sharing (Dawes, 1996; Gil-Garcia & Sayogo, 2016). Technology investments and changes in recent years might have decreased technological barriers and offered new tools that allow small cities to engage in data sharing.

Cloud-based technologies are an example of new tools that cities might utilize to share data. I find that city departments which utilize cloud-based file sharing technologies (e.g., Dropbox or Google docs) report greater data access from nongovernmental organizations. Departments in the same city might rely on internal information systems – e.g., intranet - and public agencies might have dedicated systems to share data and information that strictly comply with cybersecurity and privacy regulations and standards. Cloud-based technologies might provide a new tool for public managers to overcome economic (e.g., costs) and technical (e.g., capacity) barriers that hinder the implementation of data sharing systems with non-governmental actors. They can also help organizations to develop common data standards and address interoperability constraints.

However, these tools are often designed by private organizations, and public agencies might have little control over privacy and security issues. Public managers should pay attention to the usage of private-owned file sharing tools and be aware of the possible consequences of data breaches and unintended disclosure of information. Some file-sharing tools also have specific policies regarding property rights on the materials stored on their server (e.g., Dropbox). Public managers should ascertain that the chosen tool fits with the scope and type of collaboration that they aim to achieve.

In contrast with technological barriers, socio-political barriers are relevant to data access. Socio-political barriers include issues related to political concerns and competition that might prevent data sharing. For instance, data might be too sensitive to be shared, or local politicians might oppose the exchange of information for political reasons. Socio-political barriers have received little attention in data sharing literature and should be further integrated into the IFDA and future studies. Socio-political barriers stem from underlying issues related to institutions and power relationships between public managers as bureaucrats and politicians as representatives of civil society interests (Rainey, 2009). In the case of city governments, researchers should consider how political competition between state and local governments or between city departments and local stakeholder groups affect the exchange of information.

Results suggest that departments in principal cities of metropolitan or micropolitan areas are more likely to have access to data. Growing urbanization has created large metropolitan areas which require city governments to coordinate activities and share data and information with multiple stakeholders to provide public services, including transportation, welfare assistance, and public safety (Lefèvre, 1998; Parks & Oakerson, 1989). Cities that are main centers of metropolitan areas are more likely to have resources and influence to affect other organizations' behavior. Moreover, by coordinating activities with multiple stakeholders, principal cities might have strong relationships with other public agencies and nongovernmental organizations, which facilitate data access. While available data do not allow further investigation, future studies should concentrate on data access and exchange across local governments in the same metropolitan (or micropolitan) area to understand how environmental characteristics stemming from geographical proximity, shared policy problems, competition for legitimacy, residents, and funding, and collaborative initiatives affect data access.

Finally, the analysis shows that there are differences in data access across department types. Police departments are the most likely to access data, while community development departments report lower data access as compared to the Mayor's office. Variation in the institutional and social context in which departments operate might explain these differences. Police departments have a hierarchical structure that links them to state and federal agencies and creates a set of legal requirements that police chiefs need to address. (Grimmelikhuijsen & Feeney, 2016; Meijer & Torenvlied, 2016). A hierarchical structure might facilitate data access by establishing and enforcing centralized rules for sharing data. Moreover, police departments have the authority to force nongovernmental and public organizations to provide data to address security and safety issues. Future research should consider how specific differences across departments, such as differences in the task environments, field norms and professional rules (Yavuz & Welch, 2014) affect a local department's ability to access and share data.

Summary. The Integrative Framework for Data Access provides a comprehensive assessment of how public managers can build systemic capacity to access data from internal and external stakeholders. The IFDA combines insights from different frameworks (Bellamy et al., 2008; Fountain, 2007; Gil-Garcia et al., 2010; Yang et al., 2014) to provide a more nuanced understanding of data access. Moreover, it discusses the importance of coordination mechanisms to share data. Informal coordination and lateral coordination are especially important when public organizations have low capacity and face financial constraints because they are less costly to implement than routines and data sharing agreements. Moreover, this research compares results across three different types of stakeholders – other departments in the city, other public agencies, and nongovernmental organizations. Acknowledging commonalities and differences across relationship types is important as previous studies do not recognize the heterogeneity of participants to data sharing initiatives.

The multilevel structure of the vertical dimension shows how exogenous (not controlled by public managers) and endogenous (controlled by public managers) factors contribute to an organization's ability to access data. Results indicate that legal mandates to share data, external influence dynamics, and informal coordination mechanisms explain the most variation in data access across city departments. The horizontal dimension suggests that there is variation in data access across the portfolio of stakeholders as previous research has shown that barriers to data exchange and access vary across private, nonprofit, and public organizations (Azad & Wiggins, 1995; Dawes et al., 2009; Roberts, 2011). Results confirm this hypothesis, and table 19 summarizes common factors that characterize internal vs. external stakeholders and same sector vs. cross-sector relationships.

Table 19 shows that public managers who rely on informal coordination are more likely to obtain data from other organizations in the public sector. As suggested above, public managers might have more extensive and closer networks with managers and employees working in the same sector. Moreover, legal mandates strongly increase the responsiveness of other public organizations and facilitate data access. When taken together, these findings suggest that in same-sector relationships data access is predicted by a mix of formal institutions and informal coordination mechanisms. Future research should follow this lead and further investigate how formal institutions and informal mechanisms are balanced in same-sector relationships for sharing data and information.

When exchanging data with external stakeholders, public managers face challenges and opportunities related to power dynamics in the social environment. When exchanging data with other public agencies and nongovernmental organizations, public managers should consider the social environment in which they operate. Power dynamics greatly shape opportunities and constraints to share data, and future research should explore how managers can overcome these constraints.

Finally, when looking at cross-sector relationships, we note that coordination mechanisms are less effective and legal mandates have a lower effect on data access. Results also show that file-sharing tools have a positive impact on data access in cross-sector relationships. This result might indicate that the lack of shared infrastructure is a significant antecedent of cross-sector data access.

Table 19.

Summary results, horizontal dimension of the IFDA

	Internal	*External*
Public (same) sector	• High impact of legal mandates • Low impact of the social environment (influence) • Informal coordination mechanisms	• High impact of legal mandates • High impact of the social environment • Informal coordination mechanisms
Cross sector	NA	• Low impact of legal mandates • High impact of the social environment • Information sharing tools

Limitations

Before proceeding to implications for practice and future research, I acknowledge the limitations of the study. Some limitations are intrinsically part of survey methodology, such as common method bias, while others concern the measurement of the variables and data collection.

Common method bias. Survey data suffer from common method bias. Common method bias occurs when the data variance is to be attributed to the measurement method (e.g., repeated questions to the same individual) rather than to a natural variation in the observed social phenomena. When this occurs, the correlations across the variables are artificial and would not be observed with proper measurement (Podsakoff et al. 2003).

According to Conway and Lance (2010) researchers can apply three strategies to reduce the potential effect of common method biases in survey research: (i) justify the appropriateness of self-reported variables; (ii) provide evidence of construct validity and lack of overlap across measures; and (iii) proactively consider common method bias issues at the design stage. I discuss how I utilized each strategy.

First, self-reported measures are appropriate for several constructs in the study. For instance, Pfeffer & Salancik (1978) note that managers make decisions based on how they perceive power and influence in the social environment. Therefore, considering public managers' perceptions of stakeholder influence is an appropriate measure. Moreover, the research aims to investigate how public managers coordinate data access; in this case, public managers are the best source of information to understand how frequently they utilize each coordination mechanism.

With regards to the measures of technical capacity and technology use, research suggests that self-reported measures of technology use and capacity are often biased as high users tend to underreport hours of use while light users over report hours of use (Collopy, 1996). However, at the aggregated level over- and under-estimated measures tend to regress to the mean, reducing the overall bias. Finally, institutional variables are not subject to common method bias because they are constructed using observational data from other sources, and therefore they are not subject to self-reporting issues.

Second, whenever appropriate and possible, the study relies on constructs tested in previous literature and previous versions of the survey (e.g., technical capacity or influence items). All variables show excellent internal validity as shown by the high value of their Cronbach's alpha. When previous scales and measures were not available, I provided information on the underlying factor structure, such as in the case of Technical and Socio-Political Barriers. Results of the factor analysis are provided to the readers to evaluate the overlap across measures and internal consistency of the constructs.

Third, the research team was attentive to adopt strategies to reduce common method bias during the survey design stage, including intermixing survey items, protecting individual anonymity, and improving scale items (Podasakoff et al. 2003). Moreover, the survey included a set of ‘quality check’ items, for which respondents were explicitly asked to mark a given answer option. For instance, respondents were instructed to mark the “Strongly agree” or “Strongly disagree” option in three questions throughout the survey. Quality checks help to control for “the propensity for respondents to try to maintain consistency in their responses to questions” (Podasakoff et al., 2003, p. 884). The quality checks were utilized in the data cleaning phase to identify respondents who consistently checked the same category and did not follow the instructions provided. These respondents were eliminated from the final dataset.

Measurement limitations. This research focuses only on one dimension of data access, which is timeliness. Timely access to data is fundamental to provide public managers with the information they need to guide their decisions on public policies and services. Delayed access to data might reduce the ability of city governments to respond to stakeholders’ needs and to unforeseen situations and might increase the amount of resources (time, human, financial) that public managers need to invest in obtaining data. However, it might be that public managers will be able to access data after several requests; this aspect is not considered by current research. Internal and external stakeholders might not have the capacity to promptly respond to city departments’ requests and might need extra time to prepare and share the data. Future research should consider how the IFDA explains the overall access to data and the quality of data that public managers are able to obtain. By taking more time to respond to the city department, external organizations might be able to provide “better” data.

There are several dimensions of quality that are important for public organizations. For instance, data need to be usable, so that public managers can readily analyze them. Data shared in PDF format, for instance, require time to be translated into a format readable by statistical software. There are also issues related to data collection methods, representativeness, and accuracy of data that are of the utmost importance to design effective, equitable and fair public policies. Future research should consider how the IFDA explains variation in quality and usability of data that public managers obtain from other organizations; even when public managers can timely access data, they might not get the data they need or want to inform their decisions.

Second, the survey focuses only on public managers’ perceptions and does not collect information on the external stakeholders. Because of this, we need to be cautious when interpreting some results. For instance, influence dynamics do not consider the perspective of both actors involved. The dyadic nature of power is one of the core tenants of resource dependence theory (Emerson, 1962; Saidel, 1991). In the study, I have no information on how other organizations perceive their relative power over or dependence from the city government; influence dynamics could have been different, had both sides been taken into account. While a city department might perceive individual citizens to be very influential on policymaking, citizens might perceive that the city government makes the final decision and political actors are the major players of policymaking. Therefore, they might grant access to data because they perceive the city department to have power and control over policy decisions.

Moreover, I do not have information regarding the capacity of the stakeholders or their organizational structure. Organizations might refuse to provide data to the city government because they do not have the capacity to share them. Small nonprofit organizations, for instance, might not store data in a digital format or might not store the data that the city department needs. The Technical Barriers variable controls for these challenges, but it is self-reported by city department heads. Additionally, some organizations might have strict organizational rules and employees might lack the authority to decide to share the data. Future research should consider network data to explore both sides of data access: the data provider and the data receiver.

Finally, stakeholders are aggregated into three main categories - other departments in the city, other public agencies, and nongovernmental organizations; because of this data structure, the research cannot consider the characteristics of every single relationship. Within the same category, there is likely variation about how often a city department interacts with some actors rather than others; the influence that actors have over the policy-making process; or even coordination mechanisms that public managers apply. The research overlooks dyadic differences which should further explore in future research.

Data collection. There are two main limitations with regards to data collection. First, some states have only one or two observations as only department heads from one or two cities responded to our survey. Because of this, the coefficients and significance of the institutional variables are based on very small clusters, and they might be biased. The small number of observations per cluster also prevented cross-level interactions between level-1 and level-2 variables which this study does not explore. Researchers should further investigate to what extent the institutional context moderates the effectiveness of coordination mechanisms that public managers utilize.

Second, data are cross-sectional. Cross-sectional data do not allow to empirical test causation; causation is theoretically justified. Endogeneity issues might affect the reliability of results on coordination mechanisms applied by managers to access data. Public managers might have selected coordination mechanism based on their previous experience. If the selection of the coordination mechanisms is nonrandom, then estimation will suffer from biased results. Unfortunately, the survey instruments do not provide a good instrumental variable that can be used to estimate a two-stage regression which is usually applied to deal with endogeneity issues (Greene, 2000).

Kelman, Hong, and Turbitt (2013) provide evidence that public managers randomly select strategies, as they are not aware of which one will be the most successful – otherwise, they will use it since the beginning. If we make a similar assumption for this study, we can expect that the estimation will not be biased. Data sharing literature also suggests that the selection of coordination mechanisms to exchange data is based on the frequency of the data exchange rather than the outcome (Roberts, 2011). Therefore, we can expect that data access is not likely to be a key determinant of coordination mechanisms. Future studies might apply different methodologies, such as experimental designs, to test the effectiveness of coordination mechanisms and rule out endogeneity issues. An experimental approach has been used to test the effectiveness of Freedom of Information regulations (Worthy et al., 2017) and might be adequate to expand research on data access.

Implication for Public Management Research

This study provides practical implications for public organizations that aim to increase data access in the public sector, both within and across organizational boundaries, and public managers who want to intervene at the micro level to reduce barriers for cooperation and exchange.

This research suggests that state-level interventions might not be effective to increase data access. State government managers and policymakers might need to consider new strategies and initiatives to support city governments to access data beyond institutional support or symbolic initiatives like open data portals. For instance, results show that reducing government influence on a city department's decision-making processes might facilitate access to data from both government and nongovernmental organizations. Previous studies have noted that there is great variability in how states manage their relationships with local governments (Mullin & Daley, 2010). State and federal government agencies should consider how to promote collaborative relationships with local governments that grant autonomy and limits state government's interference in local decisions.

Among institutions, the only significant variable is legal mandate, which increases data access across the portfolio of relationships. Policymakers should consider developing and approving laws to legally mandate the provision of data to government agencies, particularly when public managers need data to address sensitive policy issues or when delays might negatively affect government activity.

At the micro-level, results show that informal coordination is still an important determinant of intra-sector data access. Public managers could increase network development activities, including frequent meetings, workshops, and interagency events for managers and employees to socialize. Moreover, if data access mostly relies on interpersonal relationships, public organizations should pay attention to internal turnover, as changes in managerial positions might affect data access. It is possible that less connected department heads will be less likely to access data; for instance, younger managers or managers that have recently transferred from other cities could encounter more significant challenges to access data. Networking activities can also be encouraged by nonprofit organizations, such as professional associations or foundations, which aim to promote data practices in the public sector. These might include professional events and workshops, online and in-person training, and collaborative initiatives.

Future Studies

While previous sections have indicated several opportunities for future research, this last section suggests four main research areas that are particularly relevant to public management scholarship and data sharing literature.

First, this study tests the IFDA looking only at three categories of stakeholders: other departments in the city; other public agencies; and nongovernmental organizations. Future studies should test the framework across other categories and typologies of stakeholders to provide a more nuanced understanding of the horizontal dimension. For instance, researchers could separate vertical relationships - e.g., federal, state, and county agencies - from horizontal relationships - e.g., other city governments. Different power dynamics characterize vertical and horizontal relationships. Collaboration studies suggest that relationships between the federal government and city governments are based on coercive power, while relationships between the state government and city governments are more collaborative and equitable (Mullin & Daley, 2010). Similarly, relationships with for-profit organizations might differ from relationships with nonprofit ones. Nonprofit organizations have greater dependence from public organizations and might be more willing to provide data (Guo & Acar, 2005). For-profit organizations are more concerned about property issues and copyright and are less willing to exchange data with other organizations (Meijer, 2018).

Second, this study does not investigate the perspective of the external stakeholders (e.g., employees and managers working outside the city or in other city departments). Previous studies have shown that mutual dependence and reciprocity might positively affect the frequency of data and information exchange (Galaskiewicz & Marsden, 1978; Including et al., 2013). I also do not have information on organizational policies and practices in providing organizations. Future studies should utilize network data to collect information at the dyadic level on both sides of the exchange and test the theoretical mechanisms proposed in this framework.

Third, with regards to coordination mechanisms, the IFDA provides little explanation for data access from nongovernmental organizations. Future research should dedicate more attention to investigate the antecedents of data access from nonprofit and for-profit organizations. Local governments are increasingly interested in accessing data collected by nonprofit and for-profit organizations; these organizations own data that are often not available to public agencies or are expensive to collect. Partnerships and synergies with nongovernmental organizations are expected to improve the quality and the availability of data and in turns decision-making processes and outcomes for public organizations.

For instance, future studies should consider the composition and structure of managers' networks. While there is a growing number of studies investigating intra-organizational networks among public employees (Nisar & Maroulis, 2017; Siciliano, 2015, 2016), we know less about inter-organizational networks. It might be that interpersonal networks are less effective in cross-sector relationships because public managers have fewer or weaker connections. Researchers could also focus on city governments that are particularly successful in exchanging data with nongovernmental organizations to understand the characteristics of the coordination mechanisms in place.

Finally, future research should consider what prompts managers to adopt each coordination mechanism and under what conditions these coordination mechanisms are more effective. Collaboration studies drawing from a contingency perspective (Scott, 2003; Thompson, 1967) have shown that the effectiveness of coordination mechanisms might vary depending upon organizational and environmental characteristics (Kelman et al., 2013). Formal coordination might be effective when the organization has financial and human resources to guarantee the implementation of routines and data sharing agreements. Lateral coordination might be more effective when organizations exchange data for the first time as public managers lack significant inter-personal relationships but become less effective over time. Understanding these contingencies will improve our theoretical and empirical understanding on how public managers can share data with internal and external stakeholders.

REFERENCES

6, P. (2004). Joined-Up Government in the Western World in Comparative Perspective: A Preliminary Literature Review and Exploration. Journal of Public Administration Research and Theory, 14(1), 103–138. https://doi.org/10.1093/jopart/muh006

6, P., Bellamy, C., Raab, C., Warren, A., & Heeney, C. (2007). Institutional Shaping of Interagency Working: Managing Tensions between Collaborative Working and Client Confidentiality. Journal of Public Administration Research and Theory, 17(3), 405–434. https://doi.org/10.1093/jopart/mul018

Adler, P. S., & Kwon, S.-W. (2009). Social Capital: The Good, the Bad, and the Ugly (SSRN Scholarly Paper No. ID 186928). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=186928

Agranoff, R., & McGuire, M. (2001). Big Questions in Public Network Management Research. Journal of Public Administration Research and Theory, 11(3), 295–326.

Aiken, M., Bacharach, S. B., & French, J. L. (1980). Organizational Structure, Work Process, and Proposal Making in Administrative Bureaucracies. The Academy of Management Journal, 23(4), 631–652.

Akbulut-Bailey, A. Y. (2011). Information Sharing Between Local and State Governments. The Journal of Computer Information Systems; Stillwater, 51(4), 53–63.

Al-Alawi, A. I., Al-Marzooqi, N. Y., & Mohammed, F. (2007). Organizational culture and knowledge sharing: critical success factors. Journal of Knowledge Management, 11(2), 22–42. https://doi.org/10.1108/13673270710738898

Allard, S. W., Wiegand, E. R., Schlecht, C., Datta, A. R., Goerge, R. M., & Weigensberg, E. (2018). State Agencies' Use of Administrative Data for Improved Practice: Needs, Challenges, and Opportunities. Public Administration Review, 78(2), 240–250. https://doi.org/10.1111/puar.12883

Alonso, J. M., Clifton, J., & Díaz-Fuentes, D. (2015). Did New Public Management Matter? An empirical analysis of the outsourcing and decentralization effects on public sector size. Public Management Review, 17(5), 643–660. https://doi.org/10.1080/14719037.2013.822532

Ansell, C., & Gash, A. (2008). Collaborative Governance in Theory and Practice. Journal of Public Administration Research and Theory, 18(4), 543–571. https://doi.org/10.1093/jopart/mum032

Anssi Smedlund. (2008). The knowledge system of a firm: social capital for explicit, tacit and potential knowledge. Journal of Knowledge Management, 12(1), 63–77. https://doi.org/10.1108/13673270810852395

Argote, L., Gruenfeld, D., Naquin, C., & Turner, M. E. (2001). Group learning in organizations. Groups at Work: Theory and Research, 369–411.

Atabakhsh, H., Larson, C., Petersen, T., Violette, C., & Chen, H. (2004). Information Sharing and Collaboration Policies within Government Agencies. In Intelligence and Security Informatics (pp. 467–475). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-25952-7_37

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. Government Information Quarterly, 32(4), 399–418. https://doi.org/10.1016/j.giq.2015.07.006

Avery, G. (2000). Outsourcing Public Health Laboratory Services: A Blueprint for Determining Whether to Privatize and How. Public Administration Review, 60(4), 330–337. https://doi.org/10.1111/0033-3352.00095

Azad, B. (1998). Management of entreprise-wide GIS implementation: Lessons from exploration of five case studies (PhD Thesis). Department of Urban Studies and Planning, MIT, Cambridge, MA.

Azad, B., & Wiggins, L. L. (1995). Dynamics of inter-organizational geographic data sharing: a conceptual framework for research. In H. J. Onsrud & G. Rushton (Eds.), Sharing Geographic Information (pp. 22–43). New Brunswick, NJ: Center for Urban Policy Research.

Bajaj, A., & Ram, S. (2007). A Comprehensive Framework Towards Information Sharing Between Government Agencies. International Journal of Electronic Government Research; Hershey, 3(2), 29–44.

Bate, S. p., & Robert, G. (2002). Knowledge management and communities of practice in the private sector: lessons for modernizing the National Health Service in England and Wales. Public Administration, 80(4), 643–663. https://doi.org/10.1111/1467-9299.00322

Bauhr, M., & Grimes, M. (2013). Indignation or Resignation: The Implications of Transparency for Societal Accountability. Governance, 27(2), 291–320. https://doi.org/10.1111/gove.12033

Baumer, D. L., Earp, J. B., & Poindexter, J. C. (2004). Internet privacy law: a comparison between the United States and the European Union. Computers & Security, 23(5), 400–412. https://doi.org/10.1016/j.cose.2003.11.001

Bekkers, V. (2007). The governance of back-office integration. Public Management Review, 9(3), 377–400. https://doi.org/10.1080/14719030701425761

Bekkers, V. (2009). Flexible information infrastructures in Dutch E-Government collaboration arrangements: Experiences and policy implications. Government Information Quarterly, 26(1), 60–68. https://doi.org/10.1016/j.giq.2007.09.010

Bell, B., Ferron, J., & Kromrey, J. (2008). Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models. Presented at the Section on Survey Research Method - JSM 2018. Retrieved from https://ww2.amstat.org/sections/srms/Proceedings/y2008/Files/300933.pdf

Bellamy, C., 6, P., Raab, C., Warren, A., & Heeney, C. (2008). Information-Sharing and Confidentiality in Social Policy: Regulating Multi-Agency Working. Public Administration, 86(3), 737–759. https://doi.org/10.1111/j.1467-9299.2008.00723.x

Berman, E. M., West, J. P., & Richter, M. N. (2002). Workplace relations: Friendship patterns and consequences (according to managers). Public Administration Review, 62(2), 217–230.

Berry, F. S., & Berry, W. D. (2014). Innovation and Diffusion Models in Policy Research. In P. A. Sabatier & C. M. Weible (Eds.), Theories of the Policy Process (pp. 307–359).

Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. Government Information Quarterly, 27(3), 264–271. https://doi.org/10.1016/j.giq.2010.03.001

Betsill, M. M., & Bulkeley, H. (2004). Transnational Networks and Global Environmental Governance: The Cities for Climate Protection Program. International Studies Quarterly, 48(2), 471–493. https://doi.org/10.1111/j.0020-8833.2004.00310.x

Bettencourt, L. M. A. (2014). The Uses of Big Data in Cities. Big Data, 2(1), 12–22. https://doi.org/10.1089/big.2013.0042

Black, L. J., Cresswell, A. M., Luna, L. F., Pardo, T. A., Martinez, I. J., Thompson, F., … Cook, M. (2003). A dynamic theory of collaboration: a structural approach to facilitating intergovernmental use of information technology. In 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the (pp. 12 pp.-). https://doi.org/10.1109/HICSS.2003.1174222

Blau, P. M. (1955). The dynamics of bureaucracy: A study of interpersonal relations in two government agencies. Chicago, U. P.

Bock, G.-W., Zmud, R. W., Kim, Y.-G., & Lee, J.-N. (2005). Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate. MIS Quarterly, 29(1), 87–111.

Bolino, M. C., Turnley, W. H., & Bloodgood, J. M. (2002). Citizenship Behavior and the Creation of Social Capital in Organizations. The Academy of Management Review, 27(4), 505–522. https://doi.org/10.2307/4134400

Bostic, B. (2015, August). Big Data and Modernizing Federal Statistics: Update. US Department of Commerce - Economics and Statistics Administration. Retrieved from https://www2.census.gov/cac/sac/meetings/2015-09/2015-Bostic.pdf

Bostrom, R. P., & Heinen, J. S. (1977). MIS Problems and Failures: A Socio-Technical Perspective, Part II: The Application of Socio-Technical Theory. MIS Quarterly, 1(4), 11–28. https://doi.org/10.2307/249019

Bozeman, B., & Bretschneider, S. (1986). Public management information systems: Theory and prescription. Public Administration Review, 475–487.

Brant, R. (1990). Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. Biometrics, 46(4), 1171–1178. https://doi.org/10.2307/2532457

Bretschneider, S. (1990). Management information systems in public and private organizations: An empirical test. Public Administration Review, 536–545.

Bryson, J. M., Crosby, B. C., & Stone, M. M. (2006). The Design and Implementation of Cross-Sector Collaborations: Propositions from the Literature. Public Administration Review; Washington, 66(S1), 44.

Burt, R. S. (2000). The Network Structure Of Social Capital. Research in Organizational Behavior, 22, 345–423. https://doi.org/10.1016/S0191-3085(00)22009-1

Bushway, S., Johnson, B. D., & Slocum, L. A. (2007). Is the Magic Still There? The Use of the Heckman Two-Step Correction for Selection Bias in Criminology. Journal of Quantitative Criminology, 23(2), 151–178. https://doi.org/10.1007/s10940-007-9024-4

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. Journal of Human Resources, 50(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2012). Robust inference with multiway clustering. Journal of Business & Economic Statistics. Retrieved from http://amstat.tandfonline.com/doi/abs/10.1198/jbes.2010.07136

Chan, C. M. L. (2013). From Open Data to Open Innovation Strategies: Creating E-Services Using Open Government Data. In 2013 46th Hawaii International Conference on System Sciences (HICSS) (pp. 1890–1899). https://doi.org/10.1109/HICSS.2013.236

Chang, H. H., & Chuang, S.-S. (2011). Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. Information & Management, 48(1), 9–18. https://doi.org/10.1016/j.im.2010.11.001

Chen, Y.-C., & Lee, J. (2018). Collaborative data networks for public service: governance, management, and performance. Public Management Review, 20(5), 672–690. https://doi.org/10.1080/14719037.2017.1305691

Chen, Y.-C., & Thurmaier, K. (2009). Interlocal Agreements as Collaborations: An Empirical Investigation of Impetuses, Norms, and Success. The American Review of Public Administration, 39(5), 536–552. https://doi.org/10.1177/0275074008324566

Chow, W. S., & Chan, L. S. (2008). Social network, social trust and shared goals in organizational knowledge sharing. Information & Management, 45(7), 458–465. https://doi.org/10.1016/j.im.2008.06.007

Clarke, A., & Margetts, H. (2014). Governments and Citizens Getting to Know Each Other? Open, Closed, and Big Data in Public Management Reform. Policy & Internet, 6(4), 393–417. https://doi.org/10.1002/1944-2866.POI377

Clarke, P., & Wheaton, B. (2007). Addressing Data Sparseness in Contextual Population Research: Using Cluster Analysis to Create Synthetic Neighborhoods. Sociological Methods & Research, 35(3), 311–351. https://doi.org/10.1177/0049124106292362

Cohen, G. (2017). Cultural Fragmentation as a Barrier to Interagency Collaboration: A Qualitative Examination of Texas Law Enforcement Officers' Perceptions. The American Review of Public Administration. https://doi.org/10.1177/0275074017744659

Collopy, Fred. 1996. Biases in Retrospective Self-Reports of Time Use: An Empirical Study of Computer Users. Management Science 42(5), 758 – 767.

Comfort, L. K. (2007). Crisis management in hindsight: Cognition, communication, coordination, and control. Public Administration Review, 67(s1), 189–197.

Conway, J. M., & Lance, C. E. (2010). What Reviewers Should Expect from Authors Regarding Common Method Bias in Organizational Research. Journal of Business and Psychology, 25(3), 325–334. http://dx.doi.org.ezproxy1.lib.asu.edu/10.1007/s10869-010-9181-6

Cuganesan, S., Hart, A., & Steele, C. (2017). Managing information sharing and stewardship for public-sector collaboration: a management control approach. Public Management Review, 19(6), 862–879. https://doi.org/10.1080/14719037.2016.1238102

Daley. (2009). Interdisciplinary Problems and Agency Boundaries: Exploring Effective Cross-Agency Collaboration. Journal of Public Administration Research and Theory, 19(3), 477–493. https://doi.org/10.1093/jopart/mun020

Data Protection and Security Solutions for State and Local Government. (n.d.).

Davenport, T. H., & Prusak, L. (1998). Working knowledge: How organizations manage what they know. Harvard Business Press.

Dawes, S. S. (1996). Interagency information sharing: Expected benefits, manageable risks. Journal of Policy Analysis and Management, 15(3), 377–394. https://doi.org/10.1002/(SICI)1520-6688(199622)15:3<377::AID-PAM3>3.0.CO;2-F

Dawes, S. S., Cresswell, A. M., & Pardo, T. A. (2009). From “Need to Know” to “Need to Share”: Tangled Problems, Information Boundaries, and the Building of Public Sector Knowledge Networks. Public Administration Review, 69(3), 392–402. https://doi.org/10.1111/j.1540-6210.2009.01987_2.x

Dawes, S. S., Gharawi, M. A., & Burke, G. B. (2012). Transnational public sector knowledge networks: Knowledge and information sharing in a multi-dimensional context. Government Information Quarterly, 29, Supplement 1, S112–S120. https://doi.org/10.1016/j.giq.2011.08.002

de Montalvo, U. W. (2003). In search of rigorous models for policy-oriented research: A behavioural approach to spatial data sharing. URISA Journal, 15(1), 19–28.

Department of the Environment, Transport and the Regions. (2000). Climate Change: The UK programme (p. 50). London.

Desouza, K. C., & Bhagwatwar, A. (2014). Technology-Enabled Participatory Platforms for Civic Engagement: The Case of U.S. Cities. Journal of Urban Technology, 21(4), 25–50. https://doi.org/10.1080/10630732.2014.954898

DiMaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. American Sociological Review, 48(2), 147–160.

Dodgson, M. and Gann, D. (2011): Technological Innovation and Complex Systems in Cities, Journal of Urban Technology, 18:3, 101-113

Drake, D. B., Steckler, N. A., & Koch, M. J. (2004). Information Sharing in and Across Government Agencies The Role and Influence of Scientist, Politician, and Bureaucrat Subcultures. Social Science Computer Review, 22(1), 67–84. https://doi.org/10.1177/0894439303259889

Drew, J. (2015). Sensitive but unclassified: Examinig the use of electronic information sharing systems by law enforcement agencies in the United States. Michigan State University, Michigan. Retrieved from ProQuest Dissertations Publishing.

Eglene, O., Dawes, S. S., & Schneider, C. A. (2007). Authority and Leadership Patterns in Public Sector Knowledge Networks. The American Review of Public Administration, 37(1), 91–113. https://doi.org/10.1177/0275074006290799

Emerson, R. M. (1962). Power-Dependence Relations. American Sociological Review, 27(1), 31–41. https://doi.org/10.2307/2089716

Emerson, R. M. (1976). Social Exchange Theory. Annual Review of Sociology, 2, 335–362.

Enders, C. K. (2010). Applied Missing Data Analysis (1 edition). New York: The Guilford Press.

Ertiö, T.-P. (2015). Participatory Apps for Urban Planning—Space for Improvement. Planning Practice & Research, 30(3), 303–321. https://doi.org/10.1080/02697459.2015.1052942

Esteve, M., Boyne, G., Sierra, V., & Ysa, T. (2013). Organizational Collaboration in the Public Sector: Do Chief Executives Make a Difference? Journal of Public Administration Research and Theory, 23(4), 927–952. https://doi.org/10.1093/jopart/mus035

Feeney, M. K., & Welch, E. W. (2014). Technology–Task Coupling Exploring Social Media Use and Managerial Perceptions of E-Government. The American Review of Public Administration, 0275074014547413. https://doi.org/10.1177/0275074014547413

Feeney, M. K., Welch, E., Zhang, F., Camarena, L., Cho, S., & Fusi, F. (2016). Data Sharing, Civic Engagement, and Technology Use in Local Government Agencies: Findings from a National Survey (SSRN Scholarly Paper No. ID 3069941). Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=3069941

Fishenden, J., & Thompson, M. (2013). Digital government, open architecture, and innovation: why public sector IT will never be the same again. Journal of Public Administration Research and Theory, 23(4), 977–1004.

Fountain, J. E. (2001). Building the Virtual State. Information technology and institutional change. Washington, DC: Brookings Institution Press.

Fountain, J. E. (2007). Challenges to organizational change: multi-level integrated information structures (MIIS). In V. Mayer-Schönberger & D. Lazer (Eds.), Governance and Information Technology: From Electronic Government to Information Government (pp. 63–93). MIT Press.

Frumkin, P., & Galaskiewicz, J. (2004). Institutional Isomorphism and Public Sector Organizations. Journal of Public Administration Research and Theory, 14(3), 283–307. https://doi.org/10.1093/jopart/muh028

Fusi, F., & Feeney, M. K. (2016). Social Media in the Workplace Information Exchange, Productivity, or Waste? The American Review of Public Administration, 0275074016675722. https://doi.org/10.1177/0275074016675722

Galaskiewicz, J., & Marsden, P. V. (1978). Interorganizational resource networks: Formal patterns of overlap. Social Science Research, 7(2), 89–107. https://doi.org/10.1016/0049-089X(78)90006-6

Galbraith, J. R. (1973). Designing complex organizations. Addison-Wesley Longman Publishing Co., Inc. Retrieved from http://dl.acm.org/citation.cfm?id=540368

Gazley, B. (2008). Beyond the Contract: The Scope and Nature of Informal Government-Nonprofit Partnerships. Public Administration Review, 68(1), 141–154.

Gil-Garcia, J. R., & Sayogo, D. S. (2016). Government inter-organizational information sharing initiatives: Understanding the main determinants of success. Government Information Quarterly, 33(3), 572–582. https://doi.org/10.1016/j.giq.2016.01.006

Gil-Garcia, J. R., Chengalur-Smith, I., & Duchessi, P. (2007). Collaborative e-Government: impediments and benefits of information-sharing projects in the public sector. European Journal of Information Systems; Basingstoke, 16(2), 121–133. http://dx.doi.org.ezproxy1.lib.asu.edu/10.1057/palgrave.ejis.3000673

Gil-Garcia, J. R., Pardo, T. A., & Nam, T. (Eds.). (2016). Smarter as the New Urban Agenda: A Comprehensive View of the 21st Century City. Springer International Publishing.

Gil-Garcia, J. R., Pardo, T., & Burke, G. B. (2010). Conceptualizing information integration in government. In H. Schnoll (Ed.), E-government: Information, Technology and Transformation (1st ed., pp. 179–202). New York: Routledge.

Gil-Garcia, J. R., Pardo, T., Baker, A., & others. (2007). Understanding Context through a Comprehensive Prototyping Experience: A Testbed Research Strategy for

Emerging Technologies. In System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on (pp. 104–104). IEEE.

Girth, A. M., Hefetz, A., Johnston, J. M., & Warner, M. E. (2012). Outsourcing Public Service Delivery: Management Responses in Noncompetitive Markets. Public Administration Review, 72(6), 887–900. https://doi.org/10.1111/j.1540-6210.2012.02596.x

Graber, D. A. (1992). Public sector communication. How organizations manage information. Washington, DC: CQ Press.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prevention Science: The Official Journal of the Society for Prevention Research, 8(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Graham, M., Hogan, B., Straumann, R. K., & Medhat, A. (2014). Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty. Annals of the Association of American Geographers, 104(4), 746–764. https://doi.org/10.1080/00045608.2014.910087

Greene, W. H. (2000). Econometric analysis 4th edition. International Edition, New Jersey: Prentice Hall. Retrieved from

Grimmelikhuijsen, S. G., & Feeney, M. K. (2016). Developing and Testing an Integrative Framework for Open Government Adoption in Local Governments. Public Administration Review, n/a-n/a. https://doi.org/10.1111/puar.12689

Grimmelikhuijsen, S. G., & Welch, E. W. (2012). Developing and Testing a Theoretical Framework for Computer-Mediated Transparency of Local Governments. Public Administration Review, 72(4), 562–571. https://doi.org/10.1111/j.1540-6210.2011.02532.x

Guo, C., & Acar, M. (2005). Understanding Collaboration Among Nonprofit Organizations: Combining Resource Dependency, Institutional, and Network Perspectives. Nonprofit and Voluntary Sector Quarterly, 34(3), 340–361. https://doi.org/10.1177/0899764005275411

Hale, K. (2011). How Information Matters: Networks and Public Policy Innovation. Washington, DC: Georgetown University Press.

Hamin, E. M., Gurran, N., & Emlinger, A. M. (2014). Barriers to Municipal Climate Adaptation: Examples From Coastal Massachusetts' Smaller Cities and Towns. Journal of the American Planning Association, 80(2), 110–122. https://doi.org/10.1080/01944363.2014.949590

Hannan, M. T., & Freeman, J. (1977). The Population Ecology of Organizations. American Journal of Sociology, 82(5), 929–964.

Hau, Y. S., Kim, B., Lee, H., & Kim, Y. G. (2013). The effects of individual motivations and social capital on employees' tacit and explicit knowledge sharing intentions. International Journal of Information Management, 33(2), 356–366. https://doi.org/10.1016/j.ijinfomgt.2012.10.009

Heckman, J. J. (1977). Sample Selection Bias As a Specification Error (with an Application to the Estimation of Labor Supply Functions) (Working Paper No. 172). National Bureau of Economic Research. https://doi.org/10.3386/w0172

Hedeker, D. (2015). Methods for Multilevel Ordinal Data in Prevention Research. Prevention Science: The Official Journal of the Society for Prevention Research, 16(7), 997–1006. https://doi.org/10.1007/s11121-014-0495-x

Hess, C., & Ostrom, E. (2003). Ideas, Artifacts, and Facilities: Information as a Common-Pool Resource. Law and Contemporary Problems, 66(1/2), 111–145.

Higgins, E., Taylor, M., Lisboa, P., & Arshad, F. (2014). Developing a data sharing framework: a case study. Transforming Government: People, Process and Policy, 8(1), 151–164. https://doi.org/10.1108/TG-02-2013-0007

Homburg, V., & Bekkers, V. (2002). The back-office of e-government (managing information domains as political economies). In Proceedings of the 35th Annual Hawaii International Conference on System Sciences (pp. 1666–1674). https://doi.org/10.1109/HICSS.2002.994077

Hood, C. (2006). Transparency in Historical Perspective. In C. Hood & D. Heald (Eds.), Transparency: The Key to Better Governance? (pp. 2–23). British Academy. https://doi.org/10.5871/bacad/9780197263839.003.0001

Huysman, M., & Wulf, V. (2006). IT to support knowledge sharing in communities, towards a social capital analysis. Journal of Information Technology, 21(1), 40–51.

Including, N. C. D. S. for A. D. T. S. T., Perrino, T., Howe, G., Sperling, A., Beardslee, W., Sandler, I., … Brown, C. H. (2013). Advancing Science Through Collaborative Data Sharing and Synthesis. Perspectives on Psychological Science, 8(4), 433–444. https://doi.org/10.1177/1745691613491579

Inkpen, A. C., & Tsang, E. W. K. (2005). Social Capital, Networks, and Knowledge Transfer. Academy of Management Review, 30(1), 146–165. https://doi.org/10.5465/AMR.2005.15281445

Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. Government Information Quarterly, 28(4), 446–456. https://doi.org/10.1016/j.giq.2011.01.004

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. Information Systems Management, 29(4), 258–268.

Jasanoff, S., & Martello, M. L. (2004). Earthly Politics: Local and Global in Environmental Governance. MIT Press.

Jennings, E. T., & Hall, J. L. (2012). Evidence-Based Practice and the Use of Information in State Agency Decision Making. Journal of Public Administration Research and Theory, 22(2), 245–266. https://doi.org/10.1093/jopart/mur040

Jennings, P. D., & Zandbergen, P. A. (1995). Ecologically Sustainable Organizations: An Institutional Approach. The Academy of Management Review, 20(4), 1015–1052. https://doi.org/10.2307/258964

Jimenez, B. S., Mossberger, K., & Wu, Y. (2012). Municipal government and the interactive web: Trends and issues for civic engagement. In A. Manoharan & M. Holzer (Eds.), E-Governance and Civic Engagement: Factors and Determinants of E-Democracy: Factors and Determinants of E-Democracy (pp. 251–272). Hershey: Information Science Reference.

Jun, K.-N., & Weare, C. (2011). Institutional Motivations in the Adoption of Innovations: The Case of E-Government. Journal of Public Administration Research and Theory, 21(3), 495–519. https://doi.org/10.1093/jopart/muq020

Kapucu, N., & Hu, Q. (2016). Understanding Multiplexity of Collaborative Emergency Management Networks. The American Review of Public Administration, 46(4), 399–417. https://doi.org/10.1177/0275074014555645

Karlsson, F., Frostenson, M., Prenkert, F., Kolkowska, E., & Helin, S. (2017). Inter-organisational information sharing in the public sector: A longitudinal case study on the reshaping of success factors. Government Information Quarterly, 34(4), 567–577. https://doi.org/10.1016/j.giq.2017.10.007

Katz, D., & Kahn, R. L. (1966). Organizations and the System Concept. In J. Shafritz, J. Ott, & Y. Jang, Classics of organization theory. Cengage Learning.

Kelman, S., Hong, S., & Turbitt, I. (2013). Are There Managerial Practices Associated with the Outcomes of an Interagency Service Delivery Collaboration? Evidence from British Crime and Disorder Reduction Partnerships. Journal of Public Administration Research and Theory, 23(3), 609–630. https://doi.org/10.1093/jopart/mus038

Kim, S., & Lee, H. (2006). The Impact of Organizational Context and Information Technology on Employee Knowledge-Sharing Capabilities. Public Administration Review, 66(3), 370–385. https://doi.org/10.1111/j.1540-6210.2006.00595.x

Kim, S., & Lee, J. (2012). E-Participation, Transparency, and Trust in Local Government. Public Administration Review, 72(6), 819–828. https://doi.org/10.1111/j.1540-6210.2012.02593.x

Kitchin, R. (2014). The real-time city? Big data and smart urbanism. GeoJournal, 79(1), 1–14.

Klompas, M., McVetta, J., Lazarus, R., Eggleston, E., Haney, G., Kruskal, B. A., … Platt, R. (2012). Integrating clinical practice and public health surveillance using electronic medical record systems. American Journal of Preventive Medicine, 42(6 Suppl 2), S154-162. https://doi.org/10.1016/j.amepre.2012.04.005

Koliba, C., Zia, A., & Lee, B. H. (2011). Governance informatics: Managing the performance of Inter-organizational governance networks. The Innovation Journal: The Public Sector Innovation Journal, 16(1), 1–26.

Koonin, S. E., & Holland, M. E. (2014). The Value of Big Data for Urban Science. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), Privacy, Big Data, and the Public Good - Framework for Engagement (pp. 137–156). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107590205.009

Koppenjan, J., & Klijn, E.-H. (2004). Managing Uncertainties in Networks. New York: Routledge.

Kornberger, M., Meyer, R. E., Brandtner, C., & Höllerer, M. A. (2017). When Bureaucracy Meets the Crowd: Studying “Open Government” in the Vienna City Administration. Organization Studies, 38(2), 179–200. https://doi.org/10.1177/0170840616655496

Landsbergen, D., & Wolken, G. (1998). Eliminating legal and policy barriers to interoperable government systems. Ohio Supercomputer Center, ECLIPS Program. Retrieved from https://www.osc.edu/files/press/releases/1998/phase_2_Recommendations.pdf

Leana, C. R., & Van Buren, H. J. (1999). Organizational social capital and employment practices. Academy of Management Review, 24(3), 538–555.

Lee, J. (2013). The Administrative Broker: Bureaucratic Politics in the Era of Prevalent Information. The American Review of Public Administration, 43(6), 690–708. https://doi.org/10.1177/0275074012455092

Lefèvre, C. (1998). Metropolitan government and governance in western countries: a critical review. International Journal of Urban and Regional Research, 22(1), 9–25. https://doi.org/10.1111/1468-2427.00120

Leuan, J. (2017). Data protection in the United States: overview. Thomson Reuters. Retrieved from https://uk.practicallaw.thomsonreuters.com/6-502-0467?originationContext=document&transitionType=DocumentItem&contextData=%28sc.Default%29&comp=pluk#co_anchor_a233354

Levin, M. A. (1991). The Information-Seeking Behavior of Local Government Officials. The American Review of Public Administration, 21(4), 271–286. https://doi.org/10.1177/027507409102100401

Levine, S., & White, P. E. (1961). Exchange as a Conceptual Framework for the Study of Interorganizational Relationships. Administrative Science Quarterly, 5(4), 583–601.

Lewis-Beck, C., & Lewis-Beck, M. S. (1980). Applied regression. An introduction. Thousand Oaks, CA: Sage University Paper.

Long, J. S. (1997). Regression Models for Categorical and Limited Dependent Variables. Sage Publications.

Long, J. S., & Freese, J. (2005). Regression Models for Categorical Dependent Variables Using Stata, Second Edition (2nd edition). College Station, Tex: Stata Press.

Lubell, M., Schneider, M., Scholz, J. T., & Mete, M. (2002). Watershed partnerships and the emergence of collective action institutions. American Journal of Political Science, 148–163.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. Methodology, 1(3), 86–92.

Machlup, F. (1983). The study of information: Interdisciplinary messages. Retrieved from http://philpapers.org/rec/MACTSO-9

Malatesta, D., & Smith, C. R. (2012). Government Contracts for Legal Services: Does a Previous Contracting Relationship Alter Accountability? State and Local Government Review, 44(2), 113–126. https://doi.org/10.1177/0160323X12446617

Malatesta, D., & Smith, C. R. (2014). Lessons from Resource Dependence Theory for Contemporary Public and Nonprofit Management. Public Administration Review, 74(1), 14–25. https://doi.org/10.1111/puar.12181

Malone, T. W., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. ACM Comput. Surv., 26(1), 87–119. https://doi.org/10.1145/174666.174668

Mandell, M., & Steelman, T. (2003). Understanding what can be accomplished through interorganizational innovations The importance of typologies, context and management strategies. Public Management Review, 5(2), 197–224. https://doi.org/10.1080/1461667032000066417

March, J. G., & Simon, H. (1958). Cognitive Limits on Rationality. In Organizations (pp. 157–192). McGraw-Hill.

Marvel, M. K., & Marvel, H. P. (2008). Government-to-Government Contracting: Stewardship, Agency, and Substitution. International Public Management Journal; Stamford, 11(2), 171–192.

McGuire, M., & Silvia, C. (2010). The Effect of Problem Severity, Managerial and Organizational Capacity, and Agency Structure on Intergovernmental Collaboration: Evidence from Local Emergency Management. Public Administration Review, 70(2), 279–288. https://doi.org/10.1111/j.1540-6210.2010.02134.x

McNeal, R. S., Tolbert, C. J., Mossberger, K., & Dotterweich, L. J. (2003a). Innovating in digital government in the American states. Social Science Quarterly, 84(1), 52–70.

Meerkerk, I. van, & Edelenbos, J. (2017). Facilitating conditions for boundary-spanning behaviour in governance networks. Public Management Review, 0(0), 1–22. https://doi.org/10.1080/14719037.2017.1302248

Mehra, P. (2009). A Socio-technical Perspective on New Media and Information Sharing in Government and Non-government Sectors in India An Empirical Investigation. Journal of Creative Communications, 4(1), 45–63. https://doi.org/10.1177/097325861000400104

Meijer, A. (2015). Government Transparency in Historical Perspective: From the Ancient Regime to Open Data in The Netherlands. International Journal of Public Administration, 38(3), 189–199. https://doi.org/10.1080/01900692.2014.934837

Meijer, A. (2018). Datapolis: A Public Governance Perspective on “Smart Cities.” Perspectives on Public Management and Governance. https://doi.org/10.1093/ppmgov/gvx017

Meijer, A. J., & Torenvlied, R. (2016). Social Media and the New Organization of Government Communications An Empirical Analysis of Twitter Usage by the Dutch Police. The American Review of Public Administration, 46(2), 143–161. https://doi.org/10.1177/0275074014551381

Meijer, A., & Thaens, M. (2018). Urban Technological Innovation: Developing and Testing a Sociotechnical Framework for Studying Smart City Projects, Urban Technological Innovation: Developing and Testing a Sociotechnical Framework for Studying Smart City Projects. Urban Affairs Review, 54(2), 363–387. https://doi.org/10.1177/1078087416670274

Meyer, J. W., & Rowan, B. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. American Journal of Sociology, 83(2), 340–363.

Molm, L. D. (1994). Dependence and Risk: Transforming the Structure of Social Exchange. Social Psychology Quarterly, 57(3), 163–176. https://doi.org/10.2307/2786874

Molm, L. D., Whitham, M. M., & Melamed, D. (2012). Forms of Exchange and Integrative Bonds Effects of History and Embeddedness. American Sociological Review, 77(1), 141–165. https://doi.org/10.1177/0003122411434610

Moon, M. J. (2002). The Evolution of E-Government among Municipalities: Rhetoric or Reality? Public Administration Review, 62(4), 424–433. https://doi.org/10.1111/0033-3352.00196

Moon, M. J., & Bretschneider, S. (2002). Does the Perception of Red Tape Constrain IT Innovativeness in Organizations? Unexpected Results from a Simultaneous Equation Model and Implications. Journal of Public Administration Research and Theory: J-PART, 12(2), 273–291. https://doi.org/10.1093/oxfordjournals.jpart.a003532

Moore, M. H. (1995). Creating public value: Strategic management in government. Harvard university press.

Mullin, M., & Daley, D. M. (2010). Working with the State: Exploring Interagency Collaboration within a Federalist System. Journal of Public Administration Research and Theory, 20(4), 757–778. https://doi.org/10.1093/jopart/mup029

Nahapiet, J., & Ghoshal, S. (1998a). Social capital, intellectual capital, and the organizational advantage. Academy of Management Review, 23(2), 242–266.

Nedović-Budić, Z., & Pinto, J. K. (1999). Interorganizational GIS: Issues and prospects. The Annals of Regional Science, 33(2), 183–195. https://doi.org/10.1007/s001680050100

Nedović-Budić, Z., & Pinto, J. K. (2000). Information sharing in an interorganizational GIS environment. Environment and Planning B: Planning and Design, 27(3), 455–474.

Newcombe, T. (2017, November). Small Towns Confront Big Cyber-Risks [blog]. Retrieved July 3, 2018, from http://www.govtech.com/security/GT-OctoberNovember-2017-Small-Towns-Confront-Big-Cyber-Risks.html

Newell, S., & Swan, J. (2000). Trust and inter-organizational networking. Human Relations, 53(10), 1287–1328.

Nisar, M. A., & Maroulis, S. (2017). Foundations of Relating: Theory and Evidence on the Formation of Street-Level Bureaucrats' Workplace Networks. Public Administration Review, n/a-n/a. https://doi.org/10.1111/puar.12719

Norris, D., Joshi, A., Laura, M., & Tim, F. (2018, May 1). Hackers Keep Targeting Cities. Officials Have No Idea How to Respond. [Article]. Retrieved July 3, 2018, from https://www.citylab.com/life/2018/05/why-cities-are-so-bad-at-cybersecurity/559334/

North, D. C. (1990). Institutions, institutional change and economic performance. Cambridge, NY: Cambridge University Press.

Oliveira, G. H. M., & Welch, E. W. (2013). Social media use in local government: Linkage of technology, task, and organizational context. Government Information Quarterly, 30(4), 397–405. https://doi.org/10.1016/j.giq.2013.05.019

Oliver, C. (1990). Determinants of interorganizational relationships: Integration and future directions. Academy of Management Review, 15(2), 241–265.

Pardo, T. A., Cresswell, A. M., Thompson, F., & Zhang, J. (2006). Knowledge sharing in cross-boundary information system development in the public sector. Information Technology and Management, 7(4), 293–313. https://doi.org/10.1007/s10799-006-0278-6

Parks, R. B., & Oakerson, R. J. (1989). Metropolitan Organization and Governance: A Local Public Economy Approach. Urban Affairs Quarterly, 25(1), 18–29. https://doi.org/10.1177/004208168902500103

Peters, B. G., & Pierre, J. (1998). Governance without Government? Rethinking Public Administration. Journal of Public Administration Research and Theory: J-PART, 8(2), 223–243.

Peterson, B., & Harrell, F. E. (1990). Partial Proportional Odds Models for Ordinal Response Variables. Journal of the Royal Statistical Society. Series C (Applied Statistics), 39(2), 205–217. https://doi.org/10.2307/2347760

Pew Charitable Trusts. (2018). How States Use Data to Inform Decisions: A national review of the use of administrative data to improve state decision-making. (p. 56).

Washington, DC. Retrieved from http://www.pewtrusts.org/-/media/assets/2018/02/dasa_how_states_use_data_report_v5.pdf

Pew Research Center. (2017). Public Trust in Government Remains Near Historic Lows as Partisan Attitudes Shift. Washington, DC. Retrieved from http://www.people-press.org/2017/05/03/public-trust-in-government-remains-near-historic-lows-as-partisan-attitudes-shift/

Pfeffer, J., & Salancik, G. R. (1978). Social control of organizations. In The external control of organizations. New York: Harper & Row.

Pfeffer, J., & Salancik, G. R. (2003). The external control of organizations: A resource dependence perspective. Stanford University Press.

Piotrowski Suzanne, Rosenbloom David H., Kang Sinah, & Ingrams Alex. (2018). Levels of Value Integration in Federal Agencies' Mission and Value Statements: Is Open Government a Performance Target of U.S. Federal Agencies? Public Administration Review, 0(0). https://doi.org/10.1111/puar.12937

Pitt, D., & Bassett, E. (2013). Collaborative Planning for Clean Energy Initiatives in Small to Mid-Sized Cities. Journal of the American Planning Association, 79(4), 280–294. https://doi.org/10.1080/01944363.2014.914846

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879.

Quick, K. S., & Feldman, M. S. (2014). Boundaries as Junctures: Collaborative Boundary Work for Building Efficient Resilience. Journal of Public Administration Research and Theory, 24(3), 673–695. https://doi.org/10.1093/jopart/mut085

Rainey, H. G. (2009). Understanding and Managing Public Organizations. John Wiley &

Sons.

Reynaers, A. M., & Grimmelikhuijsen, S. (2015). Transparency in Public–Private Partnerships: Not so Bad After All? Public Administration, 93(3), 609–626. https://doi.org/10.1111/padm.12142

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. Policy Sciences, 4(2), 155–169.

Roberts, N. C. (2011). Beyond smokestacks and silos: Open-source, Web-enabled coordination in organizations and networks. Public Administration Review, 71(5), 677–693.

Romzek, B. S., & Dubnick, M. J. (1987). Accountability in the Public Sector: Lessons from the Challenger Tragedy. Public Administration Review, 47(3), 227–238.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. Stata Journal, 11(2), 159–206.

Rothenberg, S., & Zyglidopoulos, S. C. (2007). Determinants of environmental innovation adoption in the printing industry: the importance of task environment. Business Strategy and the Environment, 16(1), 39–49. https://doi.org/10.1002/bse.441

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Introduction to Special Topic Forum: Not so Different after All: A Cross-Discipline View of Trust. The Academy of Management Review, 23(3), 393–404.

Saidel, J. R. (1991). Resource Interdependence: The Relationship Between State Agencies and Nonprofit Organizations. Public Administration Review, 51(6), 543.

Sayogo, D. S., Pardo, T. A., & Bloniarz, P. (2014). Information flows and smart disclosure of financial data: A framework for identifying challenges of cross boundary information sharing. Government Information Quarterly, 31, Supplement 1, S72–S83. https://doi.org/10.1016/j.giq.2013.12.004

Schein, E. H. (1985). Defining organizational culture. Classics of Organization Theory, 3, 490–502.

Scott, R. W. (2003). Organizations. Rational, Natural and Open Systems. (Fifth). New Jersey: Prentice-Hall.

Selznick, P. (1949). Foundations of the Theory of Organization. In J. Shafritz, J. Ott, & Y. Jang, Classics of organization theory. Cengage Learning.

Siciliano, M. D. (2015). Advice Networks in Public Organizations: The Role of Structure, Internal Competition, and Individual Attributes. Public Administration Review, 75(4), 548–559. https://doi.org/10.1111/puar.12362

Siciliano, M. D. (2016). Ignoring the Experts: Networks and Organizational Learning in the Public Sector. Journal of Public Administration Research and Theory, muw052. https://doi.org/10.1093/jopart/muw052

Simo, G., & Bies, A. L. (2007). The Role of Nonprofits in Disaster Response: An Expanded Model of Cross-Sector Collaboration. Public Administration Review, 67, 125–142. https://doi.org/10.1111/j.1540-6210.2007.00821.x

Smith, C. R. (2009). Institutional Determinants of Collaboration: An Empirical Study of County Open-Space Protection. Journal of Public Administration Research and Theory, 19(1), 1–21. https://doi.org/10.1093/jopart/mum037

Snijders, T. A. B., & Bosker, R. J. (1993). Standard Errors and Sample Sizes for Two-Level Research. Journal of Educational Statistics, 18(3), 237–259. https://doi.org/10.3102/10769986018003237

Spender, J.-C. (1996). Making knowledge the basis of a dynamic theory of the firm. Strategic Management Journal, 17(S2), 45–62.

Steinhardt, B. (2010, November 6). Weak privacy protections in the U.S. The Washington Post. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2010/11/05/AR2010110507420.html

Stone, D. (2001). Policy Paradox: The Art of Political Decision Making (Revised edition). New York: W. W. Norton & Company.

Suárez, D. F. (2011). Collaboration and Professionalization: The Contours of Public Sector Funding for Nonprofit Organizations. Journal of Public Administration Research and Theory, 21(2), 307–326. https://doi.org/10.1093/jpart/muq049

Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as “bazaars”?: A review of coordination problems and mechanisms to match demand for data with supply. Transforming Government: People, Process and Policy, 11(1), 157–172. https://doi.org/10.1108/TG-01-2017-0007

Sunlight Foundation (n.d.). Our mission. Retrieved on November, 11 2018 from https://sunlightfoundation.com/about/

Szulanski, G. (2000). The Process of Knowledge Transfer: A Diachronic Analysis of Stickiness. Organizational Behavior and Human Decision Processes, 82(1), 9–27. https://doi.org/10.1006/obhd.2000.2884

Taylor, F. W. (1912). The principles of Scientific Management. In J. Shafritz, J. Ott, & Y. Jang, Classics of organization theory. Cengage Learning.

Thompson, J. D. (1967). Organizations in action. New York: McGraw-Hill.

Thomson, A. M., & Perry, J. L. (2006). Collaboration Processes: Inside the Black Box. Public Administration Review, 66(s1), 20–32. https://doi.org/10.1111/j.1540-6210.2006.00663.x

Tolbert, P. S. (1985). Institutional Environments and Resource Dependence: Sources of Administrative Structure in Institutions of Higher Education. Administrative Science Quarterly, 30(1), 1–13.

Townsend, A. M. (2013). Smart cities: Big data, civic hackers, and the quest for a new utopia. WW Norton & Company.

Tsai, W. (2002). Social Structure of “Coopetition” Within a Multiunit Organization: Coordination, Competition, and Intraorganizational Knowledge Sharing. Organization Science, 13(2), 179–190. https://doi.org/10.1287/orsc.13.2.179.536

Tsai, W., & Ghoshal, S. (1998). Social Capital and Value Creation: The Role of Intrafirm Networks. The Academy of Management Journal, 41(4), 464–476. https://doi.org/10.2307/257085

Tulloch, D. L., & Harvey, F. (2007). When data sharing becomes institutionalized: Best practices in local government geographic information relationships. URISA-WASHINGTON DC-, 19(2), 51.

UN Habitat. (2011). UN-Habitat Annual Report 2010 (Annual report series No. 036/11E) (p. 100). Retrieved from https://unhabitat.org/un-habitat-annual-report-2010/

Veljković, N., Bogdanović-Dinić, S., & Stoimenov, L. (2014). Benchmarking open government: An open data perspective. Government Information Quarterly, 31(2), 278–290. https://doi.org/10.1016/j.giq.2013.10.011

Vest, J. R., & Issel, L. M. (2014). Factors Related to Public Health Data Sharing between Local and State Health Departments. Health Services Research, 49(1pt2), 373–391. https://doi.org/10.1111/1475-6773.12138

Walker, A. (2017). 350 mayors adopt Paris climate accord after U.S. pulls out (updated) - Curbed. Retrieved August 1, 2017, from https://www.curbed.com/2017/6/1/15726376/paris-accord-climate-change-mayors-trump

Weber, M. (1922). Bureaucracy. In J. Shafritz, J. Ott, & Y. Jang, Classics of organization theory. Cengage Learning.

Weitzman, B. C., Silver, D., & Brazill, C. (2006). Efforts to Improve Public Policy and Programs through Data Practice: Experiences in 15 Distressed American Cities. Public Administration Review, 66(3), 386–399. https://doi.org/10.1111/j.1540-6210.2006.00596.x

Welch, E. W. (2012). The relationship between transparent and participative government: A study of local governments in the United States. International Review of Administrative Sciences, 78(1), 93–115. https://doi.org/10.1177/0020852312437982

Welch, E. W., Feeney, M. K., & Park, C. H. (2016). Determinants of data sharing in U.S. city governments. Government Information Quarterly, 33(3), 393–403. https://doi.org/10.1016/j.giq.2016.07.002

West, D. W., Mayer-Schönberger, V., & Lazer, D. (2007). Global perspectives on e-government. In Governance and information technology: from electronic government to information government (pp. 17–31). MIT Press.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: Journal of the Econometric Society, 817–838.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine, 30(4), 377–399. https://doi.org/10.1002/sim.4067

Willem, A., & Buelens, M. (2007). Knowledge Sharing in Public Sector Organizations: The Effect of Organizational Characteristics on Interdepartmental Knowledge Sharing. Journal of Public Administration Research and Theory, 17(4), 581–606. https://doi.org/10.1093/jopart/mul021

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. The Stata Journal, 6(1), 58–82. http://www.stata-journal.com/article.html?article=st0097

Worthy, B., John, P., & Vannoni, M. (2017). Transparency at the Parish Pump: A Field Experiment to Measure the Effectiveness of Freedom of Information Requests in England. Journal of Public Administration Research and Theory, 27(3), 485–500. https://doi.org/10.1093/jopart/muw063

Wukich, C., Siciliano, M. D., Enia, J., & Boylan, B. (2017). The Formation of Transnational Knowledge Networks on Social Media. International Public Management Journal, 20(3), 381–408. https://doi.org/10.1080/10967494.2016.1238428

Yang, T.-M., & Maxwell, T. A. (2011). Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors. Government Information Quarterly, 28(2), 164–175.

Yang, T.-M., & Wu, Y.-J. (2016). Examining the socio-technical determinants influencing government agencies' open data publication: A study in Taiwan. Government Information Quarterly. https://doi.org/10.1016/j.giq.2016.05.003

Yang, T.-M., Pardo, T., & Wu, Y.-J. (2014). How is information shared across the boundaries of government agencies? An e-Government case study. Government Information Quarterly, 31(4), 637–652. https://doi.org/10.1016/j.giq.2014.05.002

Yavuz, N., & Welch, E. W. (2014). Factors affecting openness of local government websites: Examining the differences across planning, finance and police departments. Government Information Quarterly, 31(4), 574–583. https://doi.org/10.1016/j.giq.2014.07.004

Zellner, A. (1963). Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results. Journal of the American Statistical Association, 58(304), 977–992. https://doi.org/10.1080/01621459.1963.10480681

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly, 31(1), 17–29. https://doi.org/10.1016/j.giq.2013.04.003

APPENDIX A

PROTOCOL FOR COGNITIVE INTERVIEWS

1. Could you please broadly describe what the term data referred to you while answering to those questions?
2. Could you please explain with your own words what question 2 is asking you?
3. Were you able to find your answer among the options provided (question 3 and 6)?
4. In question 4, we ask whether other organizations are legally required to fulfill your data requests. How hard was this to answer? How sure were you of your answer?
5. Were you able to find data barriers that usually applied to your experience?
6. Does question #8 include all quality dimensions that you generally consider when evaluating data?
7. Can you provide examples for each category mentioned in question #9? Do you generally think about ‘public accessible data’ or ‘sensitive data’?
8. Are individuals / units identify in question #10 generally involved in data sharing in your department?

APPENDIX B

SELECTION BIAS AND HECKMAN MODEL

As mentioned in Chapter 5, the survey was designed to include a screening question asking respondents whether their department "obtains data generated by other organizations to do its work." Only respondents who replied "Yes" were asked other questions concerning data sharing. Nearly two-thirds (70%) of city department heads reported obtaining data from other organizations to do their work. The analysis only includes these respondents; the remaining 30% has been excluded.

Because of the survey design, we need to consider selection biases that might occur in the analysis. Selection biases occur when the research is conducted on a non-random sub-sample of the population of interest – i.e., the cities that request data from other organizations – and the equations explaining the selection process and outcome are not independent (Bushway, Johnson, & Slocum, 2007; Greene, 2000). In other words, if the process that generates the subsample (i.e., the cities requesting data) is not independent from the process that generates the outcome of interest (i.e., data access), then results might be biased.

We can test the independence assumption using a selection process model as proposed in 1977 by Heckman. The Heckman model first estimates a probit model that predicts the likelihood of being selected into the subsample; then, it uses an OLS model (or another appropriate estimator) to predict the outcome of interest. The outcome model includes a correction factor, called the inverse Mills ratio that derives from the probit selection model and accounts for the dependence across the two models. If the Mills ratio is not significant, the equations can be assumed to be independent and be estimated separately.

Table 20 shows the Heckman selection models for each of the three dependent variables used in the study, and the selection process generated through the survey design. Drawing from theory presented in Chapter 2 and 3 (Dawes, 1996; Gil-Garcia & Sayogo, 2016, among others) and previous studies conducted on the same sample of small- and medium-sized cities (see: Welch, Feeney, & Park, 2016), I hypothesized that the selection process is determined by: the technical capacity of a city department (e.g., city departments with greater technical capacity will be more likely to request data from other organizations); the social context (e.g., city departments that engage more frequently with external stakeholders will be more likely to request data from other organizations); power and resource availability (e.g., larger city departments will be more likely to request data from other organizations); and engagement into data practices (e.g., city departments more engaged in data practices will be more likely to request data from other organizations)[19]. These variables significantly explained the variation in data sharing in previous studies (Welch et al., 2016); therefore, we can reasonably expect that they would significantly affect the selection process.

[19] Variables are presented in Chapter 4. Variables not described in Chapter 4 are presented here. Participation items measure the extent to which stakeholders participate into a city departments' decision making. Civic participation includes: individual citizens, neighborhood associations, news media, interest groups, religious groups, consultants or paid experts, professional associations, and nonprofit human service organizations (mean = 3.84, s.d. = 0.71). Government participation includes: federal government, agencies, employees and official, Governor's office, and state legislators (mean = 2.88, s.d. = 0.77). City participation includes: internal department staff, other city departments, and the Mayor's office (mean = 2.38, s.d. = 0.72). Scales range from 1= Never to 5 = Very often. All scales have a Cronbach alpha greater than 0.8. Legal Goal indicates whether the city department values legal compliance and constitutional integrity over community representation and responsiveness and organizational efficiency and effectiveness. 63 % of department heads report legal compliance and constitutional integrity as their organization main goal. Open Government measures whether the city department "has an established plan to implement open e-government" and "a common vision about open government is shared among employees" in the organization. The Cronbach's alpha is 0.54, which is acceptable given that only two items compose the scale. The mean is 3.3 with a standard deviation equal to 0.68.

The Mills ratio is also not significant in all three models. This result provides evidence that the two models are independent, and we can estimate them separately. The Wald test is non-significant ($p > 0.1$) indicating that we cannot reject the null hypothesis that the errors of outcome and selection equations are uncorrelated, providing further support for separating the two models.

.

Table 20.

Heckman selection model

	Data access from departments in the city[1]				Data access from other public agencies[2]				Data access from nongovernmental agencies[3]			
	Beta	**SE**	**P-value**		**Beta**	**SE**	**P-value**		**Beta**	**SE**	**P-value**	
Outcome model												
Institutions												
Privacy Laws	0.00	0.00	0.62		0.00	0.00	0.91		0.00	0.01	0.48	
Legal Mandate	**0.21**	**0.09**	**0.02**	*	**0.19**	**0.08**	**0.02**	*	**0.14**	**0.07**	**0.06**	+
Open government	0.01	0.02	0.71		0.03	0.02	0.12		0.04	0.03	0.14	
Institutional capacity	0.02	0.05	0.65		-0.08	0.05	0.10		-0.06	0.06	0.32	
Social environment												
City Influence	0.04	0.10	0.73		-0.02	0.07	0.74		0.01	0.07	0.91	
Civil Society Influence	0.20	0.14	0.15		**0.32**	**0.09**	**0.00**	***	**0.32**	**0.11**	**0.00**	**
Government Influence	-0.15	0.11	0.19		**-0.23**	**0.08**	**0.00**	**	**-0.22**	**0.09**	**0.01**	*
Coordination mechanisms												
Formal Coordination	0.05	0.09	0.60		0.14	0.09	0.12		0.12	0.10	0.26	
Lateral Coordination	-0.03	0.09	0.72		0.08	0.11	0.45		**0.20**	**0.09**	**0.02**	*
Informal Coordination	**0.23**	**0.09**	**0.01**	**	**0.17**	**0.10**	**0.09**	+	0.08	0.11	0.44	
Control variables												
Principal City	**0.46**	**0.16**	**0.01**	**	0.28	0.17	0.10		**0.39**	**0.20**	**0.05**	+
Social Barriers	-0.23	0.15	0.13		**-0.26**	**0.12**	**0.04**	*	**-0.18**	**0.09**	**0.05**	+
Technical Barriers	0.13	0.13	0.33		0.09	0.11	0.45		0.08	0.10	0.38	
Dept Size	-0.02	0.07	0.81		-0.06	0.06	0.29		-0.02	0.05	0.66	
Technical Capacity	-0.20	0.12	0.12		-0.13	0.10	0.21		**-0.22**	**0.09**	**0.02**	*
File Sharing Technology	**0.22**	**0.14**	**0.10**	+	0.08	0.15	0.57		**0.39**	**0.14**	**0.01**	**

Form of Government	-0.11	0.13	0.42		-0.01	0.15	0.97		0.13	0.18	0.47	
Population	**-0.27**	**0.14**	**0.06**	+	-0.12	0.11	0.28		-0.21	0.19	0.28	
Community Development	-0.34	0.24	0.15		**-0.46**	**0.22**	**0.04**	*	-0.15	0.24	0.51	
Finance	-0.37	0.30	0.22		-0.36	0.28	0.20		-0.13	0.23	0.56	
Parks and Recreation	-0.40	0.26	0.13		-0.36	0.28	0.21		-0.27	0.21	0.18	
Police	0.32	0.37	0.40		**0.51**	**0.26**	**0.05**	+	**0.40**	**0.22**	**0.07**	+
Selection model												
Technical capacity	0.00	0.11	0.97		0.00	0.11	0.98		0.01	0.11	0.95	
Civic participation	0.03	0.11	0.76		0.04	0.12	0.77		0.03	0.12	0.81	
City participation	0.11	0.11	0.34		0.11	0.11	0.33		0.11	0.11	0.32	
Government participation	0.09	0.06	0.14		0.09	0.06	0.13		**0.10**	**0.06**	**0.10**	+
Population	**0.17**	**0.09**	**0.06**	+	**0.16**	**0.09**	**0.10**	+	0.16	0.10	0.10	
Legal goal	0.01	0.05	0.80		0.00	0.04	0.92		0.01	0.04	0.86	
Form of government	-0.06	0.14	0.65		-0.08	0.13	0.56		-0.07	0.13	0.59	
Open government	**0.22**	**0.13**	**0.10**	+	0.21	0.13	0.12		0.20	0.14	0.14	
Constant	**-2.98**	**1.14**	**0.01**	**	**-2.69**	**1.15**	**0.02**	*	**-2.75**	**1.21**	**0.02**	*
/cut1	-2.98	2.16	0.17		-1.43	1.65	0.38		-1.77	3.18	0.58	
/cut2	-2.14	2.20	0.33		-0.40	1.67	0.81		-0.73	3.12	0.81	
/athrho	-0.36	0.59	0.54		-0.11	0.77	0.89		0.20	1.00	0.84	
Inverse Mill Ratio	-0.34	0.52	0.66		-0.11	0.77	0.89		0.20	0.96	0.97	
Pseudo likelihood		-1726.01				-1788.57				-1726.01		
Obs.		597				597				597		
Selected (non-selected)		396(201)				396(201)				396(201)		

[1]Wald test of indep. eqns. (rho = 0): chi2(1) = 0.37   Prob > chi2 = 0.5435

[2] Wald test of indep. eqns. (rho = 0): chi2(1) = 0.02   Prob > chi2 = 0.8870

[3] Wald test of indep. eqns. (rho = 0): chi2(1) = 0.04   Prob > chi2 = 0.8389

Other inconsistencies found by comparing responses across survey items further support the hypothesis that the subsample of cities was randomly generated. Table 21 compares responses to the screening question with responses from other survey item asking "How frequently do you receive data from people in the following types of organization?" (response options: Daily; Weekly; Monthly; Yearly; Less than once a year; Never). Among department heads that report to daily receive data from other organizations (other departments in the city, other public agencies, and nongovernmental organizations), there are several that report not obtaining data from other organizations. These inconsistencies suggest that respondents might have reported that they do not obtain data for various reasons that do not affect the outcome variable, such as the length of the survey, misinterpretation of the questions, and so on.

Table 21.

Distribution of responses across the two questions about data requests

	Other departments in the city		**Other public agencies**		**Nongovernmental organizations**	
	Obtain	**Do not obtain**	**Obtain**	**Do not obtain**	**Obtain**	**Do not obtain**
Daily	289	97	140	47	129	38
Weekly	105	59	175	72	156	62
Monthly	50	28	106	50	107	53
Yearly	7	5	25	10	37	19
Less than once a year	5	5	10	10	23	15
Never	6	2	5	7	8	9

APPENDIX C

INSTITUTIONAL CAPACITY

Table 22 shows the raw data for calculating the Institutional Capacity barriers. Data are provided by the Pew Charitable Trusts. For each state, it indicates the number of laws currently implemented regarding data governance, warehouse, inventory, integration, sharing, agreement, security, and integrity applied statewide.

Table 22.

Data-related laws, by state and topic.

State	# of laws	Governance	Warehouse	Inventory	Integration	Sharing	Agreement	Security	Integrity
AL	3	1	0	1	0	1	0	0	0
AK	0	0	0	0	0	0	0	0	0
AZ	5	0	0	0	0	4	0	1	0
AR	4	2	0	0	0	1	0	1	0
CA	10	1	0	0	2	4	0	0	3
CO	10	4	1	0	0	4	1	0	0
CT	2	0	0	0	0	1	0	1	0
DE	4	0	0	0	0	4	0	0	0
DC	0	0	0	0	0	0	0	0	0
FL	6	2	0	0	0	0	1	0	3
GA	1	0	0	0	0	1	0	0	0
HI	5	1	0	0	0	1	0	1	2
ID	3	0	1	0	0	0	0	0	2
IL	2	0	0	0	0	0	1	1	0
IN	2	0	0	0	0	2	0	0	0

IA	2	0	0	0	0	1	1	0	0
KS	0	0	0	0	0	0	0	0	0
KY	2	0	0	0	0	0	0	1	1
LA	5	2	0	0	1	1	0	0	1
ME	3	2	0	0	0	0	1	0	0
MD	3	1	0	1	0	0	1	0	0
MA	1	0	0	0	0	1	0	0	0
MI	1	0	0	0	0	1	0	0	0
MN	8	1	0	0	0	3	2	2	0
MS	4	0	1	0	0	1	0	0	2
MO	0	0	0	0	0	0	0	0	0
MT	4	0	0	0	0	0	0	1	3
NE	1	0	0	0	0	1	0	0	0
NV	1	0	1	0	0	0	0	0	0
NH	2	0	0	0	0	1	0	1	0
NJ	4	1	0	0	0	1	1	1	0
NM	8	0	0	0	0	6	0	0	2
NY	4	1	0	0	0	0	0	0	3
NC	4	1	0	0	0	1	1	1	0
ND	3	1	0	0	0	0	1	0	1
OH	2	0	0	0	0	1	0	0	1
OK	2	0	0	0	0	1	0	0	1
OR	2	2	0	0	0	0	0	0	0
PA	6	4	0	0	0	1	0	0	1
RI	3	0	0	0	0	2	0	0	1
SC	1	1	0	0	0	0	0	0	0
SD	1	0	0	0	0	0	1	0	0

TN	2	0	0	0	0	1	0	0	1
TX	5	0	0	0	0	3	2	0	0
UT	4	2	0	0	0	1	0	1	0
VT	1	0	0	0	0	1	0	0	0
VA	2	1	0	0	0	1	0	0	0
WA	3	0	0	0	0	1	0	0	2
WV	0	0	0	0	0	0	0	0	0
WI	2	0	0	0	0	0	0	1	1
WY	1	0	0	0	0	0	0	1	0

APPENDIX D

INSTITUTIONAL CAPACITY

Table 23.

Open data portals – Scores assigned on quality

Dimension	Points	Criteria
COMPLETENESS		
We evaluated each state on the data collected by Open States: bills, legislators, committees, votes and events. We also took note if a state went above and beyond to provide this information and other relevant contextual information such as supporting documents, legislative journals and schedules. Points were deducted for missing data, often roll call votes.	**0**	State provides full breadth of legislative artifacts Open States collects: bills, legislators, votes, and committees.
	-1	State does not provide stand-alone roll call votes.
TIMELINESS		
Legislative information is most relevant when it happens, and many states are publishing information in real time. Unfortunately, there are also states where updates are more infrequent and showing up days after a legislative action took place. States were dinged if data took more than 48 hours to go online.	**1**	Multiple updates throughout the day, real time or as close to it as systems will allow.
	0	Site updates once or twice daily, typically at the end of the legislative day.
	-1	Updates take longer than 24 hours to appear on the site, often up to a week.
EASE OF ACCESS		
Common web technologies such as Flash or JavaScript can cause problems when reviewing legislative data. We found that the majority of sites work fairly well without JavaScript, but some received lower	**1**	Site was considered exceptionally well layed out by multiple evaluators, no issues with Javascript.

scores due to being extremely difficult to navigate, impossible to bookmark bills, and in extreme cases, completely unusable.	**0**	Site was deemed average by those that evaluated it and/or had minor Javascript dependencies.
	-1	Site was considered more difficult than average to use by members of staff or volunteers or had more severe Javascript dependencies.
	-2	Site was considered extremely difficult to use with a heavy reliance on irregular browser behavior and Javascript.
MACHINE READABILITY		
For many sites, the Open States team wrote scrapers to collect legislative information from the website code—a slow, tedious and error prone process. We collected data faster and more reliably when data was provided in a machine-readable format such as XML, JSON, CSV or via bulk downloads. If a state posted PDF image files or scanned documents, it received the lowest score possible.	**2**	Essentially all data can be found in machine-readable formats.
	1	Lots of data in machine readable format but substantial portions that still required scraping HTML.
	0	No machine readable data but standard screen scraping techniques applied.
	-1	Site had information that was much more difficult than average to collect. (Data only accessible via PDF or that required screen scraper to emulate Javascript.)
	-2	Site had information that was unaccessible to Open States due to use of scanned PDFs.
USE OF COMMONLY OWNED STANDARDS		

Because our ability to access most of a state's data is represented by the above "Machine Readability" metric, we decided to use this provision to measure how a state made their bill text available. Making text available in HTML or PDF is the norm, and was considered an acceptable commonly owned standard (PDFs are a commonly owned standard, but it would be certainly nice to see alternative options where bill text is only available via PDF). States that only make documents available in Microsoft Word or Wordperfect formats require an individual to purchase expensive software or rely on free alternatives that may not preserve the correct formatting. It is worth noting, all states except for two met the common criteria of providing HTML and/or PDF only, one state (Kansas) went above and beyond and another (Kentucky) did not even meet this threshold.	**1**	State made an effort to go above and beyond.
	0	State provided bills in PDF and/or HTML format and nothing better (plaintext, ODT, etc.).
	-1	State only provided bills in a proprietary format.
PERMANENCE		
Many states move or remove information when a new session starts, much to the dismay of citizens seeking information on old proposals and researchers that may have cited a link (e.g. https://somelegislature.gov/HB1 vs https://somelegislature.gov/2011/HB1) only to see it point to a different bill in the following session. Tim Berners-Lee, inventor of the World Wide Web, wrote an article declaring Cool URIs Don't Change and we agree. This poses a particular challenge to us since every page on OpenStates.org points to the page we collected data from, but if a state changes their site then users lose the ability to check us against the original source. Most (but not all) states are good about at least preserving bill information, but few were equally as good about preserving information about out-of-office legislators and historical committees, equally important parts of the legislative process.	**2**	All information is avaialble in a permanent location and data goes back a reasonable amount of time (a decade or so).
	1	Almost all information has a permanent location but a single data set doesn't. (Or a recent change to the site has wiped out historical links but information appears to be preservable going forward.)
	0	Legislator & committee information lacks a permanent location (such as committees and legislators) but most is acceptable.
	-1	Ability to link to old information is badly damaged and and/or there is less than a decade of historical information.

	-2	Vital information like bills or versions lack a permanent location.

Source: Sunlight Foundation. Retrieved on March, 3rd 2018. https://openstates.org/reportcard/#criteria

APPENDIX E

CORRELATION TABLE

Table 24.

Correlation table.

	Privacy Laws (1)	Legal Mandate (City departments) (2)	Legal Mandate (Public Agencies) (3)	Legal Mandate (Non-gov. agencies) (4)	Institutional capacity (5)	Open Data Portals (6)	City Influence (7)	Civil Society Influence (8)	Government Influence (9)	Formal Coordination (10)	Lateral Coordination (11)	Informal Coordination (12)	File Sharing Tech. (13)	Socio-political barriers (14)	Technological barriers (15)	Technical Capacity (16)	Principal City (17)	Finance (18)	Mayor's Office (19)	Community Development (20)	Police (21)	Parks and Recreation (22)	Population size (23)	Dept. size (24)	Form of Government (25)
1	1																								
2	-.08	1																							
3	.03	.75	1																						
4	.04	.38	.40	1																					
5	.59	.00	.07	.10	1																				
6	.41	.00	.09	.02	39	1.00																			
7	-.03	.05	.05	.08	.04	-.08	1.00																		
8	.04	.07	.09	.10	.11	.01	.45	1.00																	
9	.05	.02	.02	.13	.06	.01	.33	.34	1.00																
10	-.04	.13	.14	.08	-.03	-.03	.03	-.09	.05	1.00															
11	.04	.14	.11	.19	.03	-.06	.02	.02	.16	.14	1.00														
12	-.04	.06	-.02	-.05	-.12	-.08	.06	.01	-.05	-.04	.14	1.00													
13	.07	.05	.05	.04	-.02	.08	.12	.14	.07	.00	-.09	-.04	1.00												
14	.07	.01	.01	.07	.02	.09	.04	.16	.13	-.06	.13	.12	-.03	1.00											
15	-.04	.01	-.02	.09	-.02	.05	.09	.25	.12	-.04	.05	.02	.08	.36	1.00										
16	-.06	.03	.00	-.03	-.04	-.02	.00	-.12	-.01	.24	.00	-.06	.10	-.14	-.15	1.00									
17	-.16	-.06	-.09	-.02	-.14	-.10	.10	.17	.12	.06	-.03	.02	.03	-.02	.08	.01	1.00								

18	.01	.08	.13	.06	.05	.00	.07	-.07	.00	.11	.04	.00	-.01	.02	-.02	.01	.02	1.00							
19	-.04	.15	.14	.07	-.04	-.02	.20	.09	.00	.09	.04	-.02	.01	-.05	-.06	.10	.00	-.18	1.00						
20	.02	-.13	-.14	-.04	.04	.03	.04	.07	-.05	-.07	-.12	-.02	.05	-.11	.05	-.14	-.01	-.28	-.26	1.00					
21	-.04	.01	.00	.02	-.06	-.03	-.26	-.18	.23	.11	.15	-.07	-.17	.07	-.05	.07	.00	-.24	-.21	-.33	1.00				
22	.04	-.09	-.10	-.10	.00	.00	-.02	.09	-.19	-.21	-.09	.13	.11	.09	.06	-.02	-.01	-.22	-.20	-.30	-.25	1.00			
23	.14	.01	-.03	.05	.20	-.01	.11	.27	.09	.07	.03	.05	.03	.07	.06	-.06	.37	.04	-.01	-.02	-.07	.05	1.00		
24	-.02	.05	.06	.05	-.03	-.09	-.05	.13	.09	.11	.12	-.08	-.07	.06	-.03	.09	.22	-.19	.09	-.29	.46	-.04	.38	1.00	
25	-.25	.05	.02	.02	.33	-.04	.04	-.02	.03	-.05	.00	.08	-.11	.05	.01	-.04	.09	-.01	-.08	.04	.03	.01	.00	-.07	1.00

APPENDIX F

ORDERED LOGIT WITH ITNERACTION BETWEEN PRIVACY LAWS AND SENSITIVE DATA

Table 25.

Ordered logit model with interaction between Privacy Law and Sensitive Data.

	Data access from other departments in the city			**Data access from other public agencies**			**Data access from nongovernmental organizations**		
	Beta	**SE**		**Beta**	**SE**		**Beta**	**SE**	
Institutions									
Privacy Laws	0.02	0.01		0.00	0.01		0.01	0.01	
Sensitive Data	0.09	0.28		-0.20	0.26		0.17	0.24	
Privacy * Sensitive data	0.00	0.00		0.00	0.00		**-0.01**	**0.00**	+
Legal Mandate	**0.39**	**0.15**	*	**0.34**	**0.14**	*	**0.27**	**0.14**	*
Open government	0.02	0.04		0.04	0.03		**0.07**	**0.04**	+
Institutional capacity	0.04	0.08		-0.12	0.08		-0.13	0.10	
Social environment									
City Influence	0.08	0.17		-0.02	0.12		0.05	0.13	
Civil Society Influence	**0.35**	**0.20**	+	**0.53**	**0.18**	**	**0.51**	**0.19**	**
Government Influence	-0.22	0.19		**-0.38**	**0.15**	*	**-0.35**	**0.15**	*
Coordination mechanisms									
Formal Coordination	0.11	0.16		**0.25**	**0.14**	+	0.18	0.16	
Lateral Coordination	-0.03	0.16		0.13	0.18		**0.34**	**0.15**	*
Informal Coordination	**0.38**	**0.14**	**	**0.31**	**0.17**	+	0.17	0.17	
Control variables									
Principal City	**0.78**	**0.28**	**	0.50	0.31		**0.68**	**0.33**	*
Social Barriers	**-0.38**	**0.22**	+	**-0.42**	**0.22**	*	**-0.29**	**0.16**	+
Technical Barriers	0.19	0.21		0.11	0.21		0.10	0.16	
Dept Size	-0.01	0.10		-0.09	0.10		-0.01	0.10	
Technical Capacity	-0.34	0.22		-0.23	0.17		**-0.39**	**0.15**	**
File Sharing Technology	0.36	0.25		0.12	0.27		**0.65**	**0.26**	*
Form of Government	-0.13	0.24		-0.05	0.22		0.26	0.27	
Population	**-0.41**	**0.17**	*	-0.19	0.17		**-0.41**	**0.15**	**
Community Development	-0.56	0.41		**-0.78**	**0.40**	*	-0.23	0.41	
Finance	-0.53	0.55		-0.54	0.47		-0.07	0.42	
Parks and Recreation	-0.70	0.45		-0.62	0.44		-0.47	0.33	
Police	0.80	0.55		**1.06**	**0.45**	*	**1.03**	**0.43**	*
/cut1	-3.53	2.60		-2.51	2.36		-3.02	1.77	
/cut2	-2.06	2.56		-0.75	2.29		-1.28	1.82	

Observation	402	402	402
Log Likelihood	-835.38	-895.70	-938.44
Pseudo R2	0.09	0.08	0.08

APPENDIX G

ROBUSTNESS CHECKS: SMALL CITIES

In the sample, there are 427 respondents from small cities, i.e., cities with a population below 100,000 inhabitants. Among these, 289 cities (68%) report that they engage in data sharing, providing a large enough sample for conducting a separate analysis. Small cities might have different characteristics than medium-sized ones, including lower capacity and resources. Results from the Population Size variable in the SUR model further confirm this hypothesis, by showing that smaller cities report lower data access from other departments in the city and nongovernmental organizations. Therefore, it might be that when coefficients are free to vary, technical capacity or technical barriers will become significant for small cities, where lack of or lower resources influences an organizational capacity to access data.

I run all models on the subset of observations including only cities with populations below 100,000 inhabitants. The models include 247 observations grouped in 38 state-level clusters. Results are shown in tables 26, 27, and 28.

Overall, results do not differ from the models that include all observations. There are no substantial differences in the institutions, social environment, and coordination mechanisms variables. Among control variables, technical capacity is a significant predictor in all models, not just the data access from "non-governmental organizations" model as when using the full sample. This result suggests that smaller cities are affected by the lack of technical capacity while larger cities have enough technical capacity to access data from other organizations.

Table 26.

Small cities - Data access from internal departments, other public agencies, and non-governmental organization – Logit model with clustered standard errors.

	Data access from internal departments in the city				Data access from public agencies				Data access from nongovernmental organizations			
	Beta	Odds ratios	SE		Beta	Odds ratios	SE		Beta	Odds ratios	SE	
Institutions												
Privacy Laws	0.01	1.01	0.01		0.00	1.00	0.01		-0.01	0.99	0.01	
Legal Mandate	**0.46**	**1.58**	**0.21**	*	**0.36**	**1.43**	**0.17**	*	**0.34**	**1.41**	**0.18**	+
Open government	-0.03	0.97	0.06		0.01	1.01	0.03		0.05	1.05	0.05	
Institutional capacity	0.07	1.07	0.10		-0.09	0.92	0.09		-0.11	0.89	0.11	
Social environment												
City Influence	0.10	1.10	0.18		0.06	1.06	0.15		0.06	1.07	0.15	
Civil Society Influence	**0.49**	**1.62**	**0.27**	+	**0.61**	**1.85**	**0.24**	**	**0.72**	**2.06**	**0.26**	**
Government Influence	-0.26	0.77	0.23		**-0.41**	**0.66**	**0.17**	*	**-0.47**	**0.63**	**0.18**	**
Coordination mechanisms												
Formal Coordination	0.09	1.10	0.21		0.25	1.29	0.21		0.16	1.17	0.25	
Lateral Coordination	-0.03	0.97	0.20		0.12	1.13	0.24		**0.37**	**1.44**	**0.20**	+
Informal Coordination	**0.41**	**1.50**	**0.17**	*	**0.32**	**1.38**	**0.19**	+	0.19	1.20	0.20	
Control variables												
Principal City	**1.07**	**2.92**	**0.41**	**	0.59	1.80	0.41		**0.74**	**2.09**	**0.42**	+
Social Barriers	**-0.52**	**0.59**	**0.28**	+	**-0.56**	**0.57**	**0.26**	*	**-0.37**	**0.69**	**0.20**	+
Technical Barriers	0.22	1.24	0.26		0.09	1.09	0.24		0.01	1.01	0.19	
Dept Size	0.01	1.01	0.14		-0.08	0.93	0.12		-0.01	0.99	0.11	

Technical Capacity	**-0.45**	**0.64**	**0.26**	+	**-0.31**	**0.73**	**0.18**	+	**-0.56**	**0.57**	**0.17**	***
File Sharing Technology	0.32	1.37	0.32		0.11	1.12	0.30		**0.78**	**2.18**	**0.33**	*
Form of Government	-0.20	0.82	0.27		0.04	1.04	0.27		0.41	1.51	0.34	
Population	0.03	1.03	0.56		0.23	1.26	0.46		-0.17	0.85	0.31	
Community Development	-0.55	0.58	0.50		-0.78	0.46	0.49		-0.24	0.78	0.47	
Finance	-0.59	0.55	0.71		-0.44	0.64	0.58		0.07	1.07	0.50	
Parks and Recreation	-0.78	0.46	0.55		-0.55	0.58	0.54		-0.51	0.60	0.39	
Police	0.37	1.45	0.68		0.86	2.37	0.53		**0.76**	**2.14**	**0.41**	+
/cut1	0.66	1.94			2.39	5.37			-1.57	3.03		
/cut2	2.16	8.65			4.14	5.30			0.28	2.95		
Obs		247				249				248		
Clusters		38				38				38		
Log pseudo-likelihood		-672.4148				-737.2568				-763.3243		
Null model log likelihood		-751.6933				-813.9027				-849.3615		
Pseudo R2		0.11				0.09				0.1		
AIC		1392.83				1522.514				1574.649		
BIC		1477.055				1606.932				1658.971		

Table 27.

Small cities - Data access from internal departments, other public agencies, and non-governmental organization – Multilevel model.

	Data access from internal departments				Data access from public agencies				Data access from nongovernmental organizations			
	Beta	**Odds ratio**	**SE**		**Beta**	**Odds ratio**	**SE**		**Beta**	**Odds ratio**	**SE**	
Institutions												
Privacy Laws	0.01	1.01	0.02		0.00	1.00	0.01		-0.01	0.99	0.02	
Legal Mandate	**0.43**	**1.54**	**0.23**	+	**0.38**	**1.47**	**0.20**	+	**0.32**	**1.38**	**0.18**	+
Institutional capacity	0.02	1.02	0.09		0.08	1.08	0.10		0.20	1.22	0.12	
Open Data Portal	0.15	1.17	0.14		-0.14	0.87	0.18		-0.09	0.91	0.18	
Social environment												
City Influence	0.23	1.26	0.21		0.07	1.07	0.16		0.02	1.02	0.17	
Civil Society Influence	0.37	1.44	0.30		**0.62**	**1.86**	**0.28**	*	**0.65**	**1.91**	**0.30**	*
Government Influence	-0.36	0.70	0.26		**-0.46**	**0.63**	**0.18**	**	**-0.50**	**0.61**	**0.19**	**
Coordination mechanisms												
Formal Coordination	0.13	1.14	0.25		0.26	1.30	0.23		0.21	1.24	0.26	
Lateral Coordination	-0.10	0.90	0.23		0.05	1.05	0.26		0.26	1.30	0.25	
Informal Coordination	**0.48**	**1.62**	**0.20**	*	**0.40**	**1.49**	**0.21**	*	**0.34**	**1.41**	**0.19**	+
Control variables												
Principal City	**1.28**	**3.60**	**0.53**	*	0.79	2.20	0.49		**0.78**	**2.18**	**0.47**	+
Social Barriers	-0.48	0.62	0.32		**-0.54**	**0.58**	**0.29**	+	-0.32	0.72	0.24	
Technical Barriers	0.23	1.26	0.29		0.13	1.14	0.26		0.09	1.09	0.22	
Dept Size	0.04	1.04	0.16		-0.04	0.96	0.15		-0.01	0.99	0.14	

Technical Capacity	**-0.51**	**0.60**	**0.29**	+	-0.30	0.74	0.21		**-0.49**	**0.61**	**0.18**	**
File Sharing Technology	0.38	1.47	0.32		0.19	1.21	0.33		**1.05**	**2.86**	**0.35**	**
Form of Government	-0.20	0.82	0.33		-0.01	0.99	0.32		0.38	1.46	0.40	
Population	-0.16	0.85	0.74		0.05	1.05	0.56		-0.37	0.69	0.40	
Community Development	-0.67	0.51	0.55		-0.88	0.41	0.59		-0.18	0.84	0.58	
Finance	-0.77	0.46	0.75		-0.50	0.61	0.67		0.08	1.08	0.62	
Parks and Recreation	-0.93	0.39	0.63		-0.57	0.57	0.62		-0.60	0.55	0.46	
Police	0.33	1.39	0.74		0.92	2.50	0.63		**0.81**	**2.25**	**0.46**	+
/cut1	-0.93	0.40	8.33		0.83	2.29	6.58		-2.70	0.07	4.03	
/cut2	0.68	1.96	8.20		2.70	14.94	6.51		-0.63	0.53	3.93	
State	**0.87**	**2.40**	**0.93**	*	**1.27**	**3.57**	**1.35**	*	**1.95**	**7.05**	**1.34**	*
Obs.		247				249.00				248.00		
Clusters		38				38.00				38.00		
Averaged cluster		6.5				6.50				6.50		
Log likelihood		-665.10132				-733.2893				-740.9		
AIC		1380.203				1516.579				1531.8		
BIC		1467.937				1604.515				1619.636		

Table 28.

Small cities - Data access from city departments, other public agencies, and non-governmental organization – SUR system

	Data access from internal capacity				**Data access from public agencies**				**Data access from nongovernmental organizations**			
	Beta	**Odds ratios**	**SE**		**Beta**	**Odds ratios**	**SE**		**Beta**	**Odds ratios**	**SE**	
Institutions												
Privacy Laws	0.00	1.00	0.01		0.00	1.00	0.00		0.00	1.00	0.01	
Legal Mandate	**0.17**	**1.18**	**0.07**	*	**0.12**	**1.13**	**0.05**	**	**0.18**	**1.20**	**0.08**	*
Institutional capacity	-0.02	0.98	0.03		0.01	1.01	0.02		0.01	1.01	0.03	
Open Data Portal	0.04	1.04	0.05		-0.07	0.93	0.06		-0.04	0.96	0.07	
Social environment												
City Influence	0.06	1.06	0.12		0.02	1.02	0.09		0.04	1.04	0.09	
Civil Society Influence	**0.36**	**1.44**	**0.16**	*	**0.47**	**1.60**	**0.12**	***	**0.44**	**1.55**	**0.13**	***
Government Influence	**-0.20**	**0.82**	**0.12**	+	**-0.27**	**0.76**	**0.10**	**	**-0.28**	**0.76**	**0.10**	**
Coordination mechanisms												
Formal Coordination	0.09	1.09	0.11		**0.19**	**1.21**	**0.11**	+	0.12	1.12	0.14	
Lateral Coordination	-0.04	0.97	0.11		0.08	1.08	0.13		**0.21**	**1.23**	**0.12**	+
Informal Coordination	**0.26**	**1.29**	**0.11**	*	**0.19**	**1.20**	**0.10**	+	0.11	1.12	0.11	
Control variables												
Principal City	**0.66**	**1.94**	**0.22**	**	0.35	1.41	0.23		**0.45**	**1.56**	**0.25**	+
Social Barriers	**-0.30**	**0.74**	**0.16**	+	**-0.32**	**0.73**	**0.14**	*	**-0.23**	**0.80**	**0.12**	+
Technical Barriers	0.12	1.13	0.16		0.01	1.01	0.14		0.00	1.00	0.10	
Dept Size	0.00	1.00	0.07		-0.04	0.96	0.07		-0.01	0.99	0.07	
Technical Capacity	**-0.26**	**0.77**	**0.13**	*	**-0.20**	**0.82**	**0.10**	+	**-0.33**	**0.72**	**0.09**	***
File Sharing Technology	0.14	1.16	0.17		0.05	1.05	0.17		**0.41**	**1.50**	**0.19**	*

Form of Government	-0.14	0.87	0.16		0.06	1.06	0.15		0.21	1.23	0.19	
Population	-0.07	0.93	0.32		0.03	1.03	0.26		-0.03	0.97	0.19	
Community Development	-0.34	0.71	0.26		**-0.51**	**0.60**	**0.24**	*	-0.14	0.87	0.28	
Finance	-0.24	0.79	0.36		-0.15	0.86	0.30		0.08	1.08	0.31	
Parks and Recreation	-0.45	0.64	0.27	+	-0.27	0.76	0.30		-0.27	0.76	0.23	
Police	**0.27**	**1.31**	**0.32**		**0.58**	**1.78**	**0.29**	*	**0.48**	**1.62**	**0.24**	*
cut1	-0.70	0.50	3.64		0.24	1.27	2.97		-0.15	0.86	1.87	
cut2	0.13	1.13	3.58		1.19	3.27	2.94		0.92	2.51	1.84	

	Beta	SE
/atanhrho_12	1.39	0.13
/atanhrho_13	0.88	0.11
/atanhrho_23	1.22	0.11

	Beta	SE
rho_12	0.88	0.03
rho_13	0.71	0.06
rho_23	0.84	0.03

APPENDIX H

ROBUSTNESS CHECKS: MULTIPLE IMPUTATION

Multiple imputation is used to impute missing data in the original dataset. Table 29 shows the amount of missing data for each dependent and independent variable included in the analysis. Variables related to coordination mechanisms - particularly lateral and informal coordination - socio-political and technical barriers, and file sharing technologies score the highest percentage of missing data (approximately from 4.5% to 5.2%). On average, the percentage of missing data is below 4%.

Table 29.

Missing data per variable in the analysis.

Variable	**Missing values**	**% missing values**	**Obs**	**Unique values**
Data access from internal departments	17	3.67%	446	3
Data access from other public agencies	17	3.67%	446	3
Data access from nongovernmental organizations	17	3.67%	446	3
Legal mandate (city departments)	16	3.46%	447	4
Legal mandate (other public agencies)	18	3.89%	445	4
Legal mandate (nongovernmental organizations)	16	3.46%	447	4
Formal coordination	5	1.08%	458	5
Lateral coordination	24	5.18%	439	5
Informal coordination	23	4.97%	440	5
Socio-political barriers	28	6.05%	435	21
Technical barriers	26	5.62%	437	18
File sharing technologies	19	4.10%	444	2
Department size	16	3.46%	447	158
Technical capacity	1	0.22%	462	21

Table 30 shows the missing data pattern. 87% of the observations contain no missing values (pattern: 111). There is a total of 28 missing data patterns in the dataset. There is no pattern that is common across a clear majority of observations.

Table 30.

Missing data patterns.

Percent	1	2	3	4	5	6	7	8	9	10	11	12	13	14
87%	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	0	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	0	0
<1	1	0	0	0	0	0	0	0	0	1	0	0	0	0
<1	1	1	0	1	1	1	1	1	1	1	0	0	0	0
<1	1	1	1	0	0	0	0	0	0	1	0	0	0	0
<1	1	1	1	0	0	1	1	1	0	1	1	1	1	1
<1	1	1	1	1	1	0	0	0	1	1	1	1	1	1
<1	1	1	0	1	1	1	1	1	1	1	0	0	1	1
<1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
<1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
<1	1	1	1	1	1	1	1	1	1	1	0	0	1	1
<1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
<1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
<1	1	0	0	1	1	1	1	1	0	1	1	0	1	1
<1	1	1	0	0	0	0	0	0	0	1	0	0	0	0
<1	1	1	0	1	1	0	0	0	1	1	0	0	0	0
<1	1	1	0	1	1	1	1	1	1	0	0	0	0	0
<1	1	1	1	0	0	1	1	1	0	0	1	1	0	0
<1	1	1	1	0	0	1	1	1	1	1	1	1	1	1
<1	1	1	1	0	1	0	0	0	0	1	1	1	1	1
<1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
<1	1	1	1	1	0	0	0	0	0	0	1	1	1	1
<1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
<1	1	1	1	1	1	0	0	0	1	1	0	0	1	1
<1	1	1	1	1	1	0	0	0	1	1	1	1	0	0

Unfortunately, no auxiliary variables are available to predict missing data. Auxiliary variables are variables that are highly correlated with variables with missing data ($r > 0.4$) or are associated with missingness patterns (Enders, 2010). The variables with missing data are specific to data access questions and were asked in a separated section of the survey (see: Chapter 5). I have tested correlation with potential auxiliary variables, such as frequency of receiving data, but the correlation is too low to help imputation (on average, 0.15). Some of the variables in model are correlated among each other (e.g., legal mandate variables or dependent variables) which improve imputation.

Multiple imputation was conducted in Stata. The imputation process used chained equations (MICE) and a separate conditional distribution for each variable imputed in the model. MICE is adapted in this case because the model includes categorical variables (White, Royston, & Wood, 2011). I created twenty-five imputed datasets. Researchers suggest to impute at least as many datasets as the average percentage of missing values (Graham, Olchowski, & Gilreath, 2007); given that the average percentage is 4%, twenty-five datasets are sufficient to reach an acceptable estimation of missing data. Moreover, I adopted a conservative approach where all variables are imputed, but imputed values are used only for the independent variables, not the dependent ones (White et al., 2011). Results from the logit model after imputation are presented in Table 31

Table 31.

Logit models with imputed independent variables

	Data access from internal departments			Data access from public agencies			Data access from nongovernmental organizations		
	Beta	**SE**		**Beta**	**SE**		**Beta**	**SE**	
Institutions									
Privacy Laws	0.00	0.01		-0.01	0.01		**-0.01**	**0.01**	+
Legal Mandate	**0.41**	**0.13**	***	**0.40**	**0.13**	**	**0.29**	**0.14**	*
Institutional capacity	0.01	0.04		0.03	0.03		0.06	0.04	
Open government	0.03	0.07		-0.10	0.08		-0.11	0.10	
Social environment									
City Influence	0.12	0.17		-0.02	0.11		0.04	0.12	
Civil Society Influence	**0.27**	**0.15**	+	**0.45**	**0.14**	***	**0.46**	**0.17**	**
Government Influence	-0.18	0.15		**-0.28**	**0.12**	*	**-0.26**	**0.11**	*
Coordination mechanisms									
Formal Coordination	0.03	0.16		0.12	0.15		0.09	0.18	
Lateral Coordination	-0.13	0.15		0.05	0.17		0.20	0.15	
Informal Coordination	**0.47**	**0.13**	***	**0.41**	**0.15**	**	**0.27**	**0.16**	+
File Sharing Technology	0.28	0.24		0.16	0.26		**0.55**	**0.23**	*
Control variables									
Principal City	**0.61**	**0.30**	*	0.45	0.29		0.48	0.32	
Social Barriers	**-0.38**	**0.21**	+	**-0.46**	**0.21**	*	**-0.27**	**0.16**	+
Technical Barriers	0.29	0.18		0.13	0.17		0.11	0.14	
Dept Size	-0.06	0.08		-0.14	0.09		-0.06	0.09	
Technical Capacity	-0.25	0.20		-0.16	0.16		**-0.31**	**0.13**	*
Form of Government	-0.20	0.24		-0.13	0.23		0.20	0.26	
Population	-0.20	0.15		-0.03	0.15		**-0.25**	**0.13**	+

Community Development	-0.57	0.38		**-0.74**	**0.39**	+	-0.35	0.41	
Finance	-0.56	0.48		-0.58	0.42		-0.25	0.38	
Parks and Recreation	-0.71	0.43		-0.69	0.42		**-0.62**	**0.32**	+
Police	0.73	0.45		**1.01**	**0.43**	*	**0.73**	**0.34**	*
/cut1	-1.72	2.31		-0.68	1.76		-2.55	1.62	
/cut2	-0.36	2.30		0.96	1.75		-0.90	1.67	
Obs.	445			445			445		
Average RVI	0.0524			0.0427			0.039		
Largest FMI	0.0518			0.043			0.0775		