City Research Online

City, University of London Institutional Repository

PUBLIC CAPACITY, PLURAL FORMS OF COLLABORATION, AND THE PERFORMANCE OF PUBLIC INITIATIVES: A CONFIGURATIONAL APPROACH[*]

Sergio G. Lazzarini
300 Rua Quatá, São Paulo, SP, Brazil. 04546-042
Insper
Sergiogl1@insper.edu.br

Leandro Simões Pongeluppe
105 St George St, Toronto, ON, Canada M5S 3E6
Rotman School of Management at the University of Toronto
L.Pongeluppe@rotman.utoronto.ca

Nobuiuki Costa Ito
2356 Alameda Santos, São Paulo, SP, Brazil. 01419-002
Ibmec São Paulo
nobuiuki.ito@ibmec.edu.br

Felippe de Medeiros Oliveira
4th floor, 106 Bunhill Row, London EC1Y 8TZ, UK
Cass Business School – City, University of London
Felippe.oliveira@cass.city.ac.uk

Armen Ovanessoff
Flat 3, 99 Crawford Street, London W1H 2HN, UK
Accenture
armen.ovanessoff@accenture.com

PUBLIC CAPACITY, PLURAL FORMS OF COLLABORATION, AND THE PERFORMANCE OF PUBLIC INITIATIVES: A CONFIGURATIONAL APPROACH

Abstract

We assess conditions that explain plural forms of public and private action using a comparative study of 24 public initiatives in Brazil, India, and South Africa. Measuring performance as evidence of positive outcomes to their target populations, we compare cases of high and low performance. Our configurational approach examines combinations of conditions leading to positive outcomes: public operational capacity, diverse collaborations nurtured by public units (with for-profit firms, with nonprofit organizations, and with other units in the pubic bureaucracy), and stakeholder orientation (permeability to multiple sources of input to design and adjust the project). We apply fuzzy set qualitative comparative analysis to unveil configurations consistent with high performance. Our configurational analysis reveals two distinct paths to high performance. A path with higher private engagement involves concurrent collaborations with for-profit and nonprofit actors, while an alternative path with higher internal (public) engagement relies on collaborations within the public bureaucracy complemented by high permeability to inputs from multiple stakeholders. Our results also confirm that strong public capacity is necessary in all high-performance configurations and support a multiple-actor, multiple-logic view of public performance.

INTRODUCTION

Many countries face a deficit of important public services for a substantial portion of their population, leading to uneven and low-quality provision of key activities in the public interest—such as transportation, water distribution, security, and education (Collier and Dollar 2002; Deaton 2013; Inoue, 2019). Addressing these voids, scholars argue, requires moving towards a more encompassing view of public action describing how multiple actors, public or private, can effectively deliver positive social outcomes (Bryson, Crosby, and Bloomberg 2015; Moore 1995; Mahoney, McGahan, and Pitelis 2009). Along these lines, a large literature has examined the *externalization* of public activities to external actors (Alford and O'Flynn 2012). Private organizations, either for-profit or nonprofit, are said to bring complementary resources and capabilities beyond what governments, alone, can deliver (Cabral, Lazzarini, and Azevedo 2013; Kivleniece and Quelin 2012; Rangan, Samii, and Van Wassenhove 2006; Brown and Potoski 2003). Some scholars even claim that, when governments fail, private action can be an

1

effective substitute for poor public services. Private engagement, in this view, helps address a deficit of essential activities not sufficiently provided by the state (Krasner and Risse 2014). Auerswald (2009, 54) goes as far as to argue that "it is precisely the failure of governments that creates opportunities for social entrepreneurs."

We however maintain that there is no single path to high performance and that the outcomes of public initiatives result from *plural* forms of action combining the interdependent effort of both public and private actors (Lynn Jr., Heinrich, and Hill 2000; Osborne 2006). We thus follow a multiple-actor, multiple-logic perspective to assess organizational paths leading to high performance (e.g. Bryson, Sancino, Benington and Sørensen, 2017). To unveil these plural paths, we adopt a configurational approach: combinations of conditions that are consistent with positive organizational outcomes (Fiss 2011; Ragin 2006). Instead of focusing on a single type of actor or collaboration, or on its marginal contribution, we unveil multiple combinations of key theoretically informed conditions that characterize public projects and, collectively, explain their positive outcomes. Thus, we can access how distinct configurations can lead to similar performance outcomes, and the underlying conditions behind each configuration.

The first condition is *public operational capacity* (Moore 1995). Research has confirmed the importance of active and capable governments to not only enforce stable rules in the public interest but also propose well-crafted projects inviting private participation (Engel, Fischer, and Galetovic 2014). In this perspective, externalization cannot generate permanent transformations if not accompanied by well-governed public bodies setting appropriate policies and mechanisms to meet societal goals (e.g. Deaton 2013). Another important condition involves the presence of multiple forms of *collaboration*. We often observe that multiple forms of collaboration are not only common but also used in tandem (Andrews and Entwitle 2010)—including external collaborations that public units forge with for-profit and nonprofit organizations (Alexander and Nank, 2009; Kivleniece and Quelin 2012) and internal collaborations between units of the public bureaucracy (Cabral and Krane Forthcoming). Yet, even without these formal collaborations, public managers may be permeable to new input as a way to improve their proposed solutions

(Denhardt and Campbell 2006; Pandey and Scott 2002). We thus consider an additional and final condition: *stakeholder orientation*, defined as the extent to which public managers create communication channels to gather input from multiple actors influencing or influenced by the project.

We analyze how these conditions lead to high performance plural configurations using a multiple-case comparative study of 24 public initiatives in Brazil, India, and South Africa. Our focus on these large economies is due to their relevant deficit of important public services for a substantial portion of their population, which has been addressed via multiple initiatives with heterogenous outcomes. We apply *fuzzy set qualitative comparative analysis* (fsQCA) to explore the relationship between different configurations of organizational conditions and performance. Starting from the coding of conditions, fsQCA applies fuzzy-set algebra to arrive at multiple combinations of conditions consistent with positive outcomes (Fiss, 2007). Following previous work (Andrews, Beynon, and McDermott 2015; Crilly 2011; Dwivedi, Joshi, and Misangyi 2018), we then use our fsQCA results to pursue theory *elaboration*. Namely, we reassess preliminary theoretical conditions proposed by extant literature in light of the results from our comparative cases and offer a refined theory of how alternative organizational paths can be associated with high performance initiatives.

Consistent with a multiple-actor, multiple-logic view of the determinants of public performance (Bryson et al. 2017; Osborne 2006), we unveil two distinct paths and explain their synergistic combination of conditions. A path with higher private engagement involves concurrent collaborations with for-profit and nonprofit actors. In this path, for-profit firms can bring novel resources and capabilities to successfully design and execute public projects, while the risk that these for-profit actors will over-emphasize the appropriation of economic value can be tempered by the simultaneous engagement of socially-oriented nonprofits. An alternative path is based on higher public engagement: high performance can occur even without external collaborations with private actors, as long as public units nurture internal collaborations within the public bureaucracy and exhibit high stakeholder orientation.

In both paths, our results confirm that strong public operational capacity is necessary to lead the project and orchestrate effective collaborations with multiple actors. An important implication is that externalization and multiple forms of collaboration are not substitutes for weak governments; when it comes to explaining high performance, research should pay attention to processes and capabilities *inside* governments. In this sense, our configurational perspective contributes to the literature by operationalizing a multiple-actor, multiple-logic perspective describing alternative paths to high performance. We also explain how these paths encompass complex performance-enhancing dynamics emanating from superior public capacity and varied forms of interaction with external actors—via either formal collaborations or a general orientation to incorporate input from outside the public sector.

THEORETICAL BACKGROUND: PUBLIC CAPACITY, COLLABORATION, AND THE PERFORMANCE OF PUBLIC INITIATIVES

In this section, we draw from the extant literature to unveil key conditions that can explain the performance of public initiatives. We start with a discussion of how to conceptualize performance, and proceed with an examination of key conditions that can explain high performance drawing from previous discussions in the public administration and general management literatures.

The Performance of Public Initiatives

While early research aligned with the New Public Management (NPM) paradigm focused on how to increase the efficiency of public bureaucracies, a more recent literature generally described as Public Value Management (PVM) has considered how multiple actors create public value not only in terms of efficient operations but also with an emphasis on outcomes valued by beneficiaries. For instance, Moore's (1995) proposed strategic management approach considers that public bureaucracies are expected to deliver services in such a way that their perceived benefits justify government action. Instead of focusing on outputs (products or services) delivered by public units, PVM conceptualizes high performance as evidence of positive *outcomes*, that is, "impacts upon those who enjoy the value/good in question or upon states of

4

nature important to those people" (Alford and O'Flynn 2009, 175). Accordingly, we define high performance as instances where the public initiative generates sustained positive outcomes valued by their target populations.

In addition, in line with the PVM literature, we also adopt a *plural* approach to explain performance: instead of assuming a marked division between public and private action, we consider a more encompassing collection of multiple actors trying to generate positive outcomes (Bryson et al. 2017; Osborne 2006; Stoker 2006). In this perspective, we next discuss a host of theoretically informed conditions that can characterize plural paths to high performance: public capacity, diverse forms of collaboration (with for-profit firms, with nonprofit organizations, and with other units in the pubic bureaucracy), and stakeholder orientation.

Public Capacity

Scholars have highlighted that the performance of public initiatives depends on the presence of a government that "transparently and efficiently serves the needs of its clients—the citizens of the state" (Fukuyama 2004, 26). In this sense, public capacity has been conceived as a government's ability to implement and execute activities promoting social development (Andrews, Pritchett, and Woolcock 2017; Geddes 1994). Recognizing the key role of public capacity in the delivery of public services, research has delved into the organizational traits that promote the performance of public execution.

A critical trait is *leadership* (Hennessey 1998). Acting as change agents, project leaders connect ideas, personnel, and resources centered on a well-defined vision of not only what needs to be done (Considine, Lewis, and Alexander 2009) but also *how* it can be done—that is, how to implement decision-making and monitoring mechanisms to promote execution and transparency (Andrews et al. 2017). Along these lines, scholars have increasingly recognized that public organizations are comprised of a set of complementary resources and managerial practices leading to effective *execution*. Thus, capable public units are expected to recruit highly skilled and professional staff (Miller and Whitford 2016); adopt mechanisms to monitor performance and curb corruption (Barzelay and Armajani 1992); and incentivize their personnel to implement

5

novel processes and promote the required adaptions (Brown and Osborne 2012). All these characteristics enhance the capacity of public units to improve project performance.

External Collaborations with For-Profit and Nonprofit Organizations

Beyond the analysis of public bureaucracies, new developments in public management have also adopted a more plural approach: instead of considering public and private action as being independent or at arm's length of one another, research has shed light on models characterized by collections of multiple parties trying to create and sustain positive societal outcomes (Brown and Potoski 2003; Bryson et al. 2015). Osborne (2006) even refers to "a *plural state*, where multiple inter-dependent actors contribute to the delivery of public services" (p. 384, emphasis in the original).

A critical issue in this discussion is how public actors can successfully engage organizations from outside the public bureaucracy, a process that has been termed *externalization* (Alford and O'Flynn 2012). Indeed, public-private interactions have been extensively studied in the literature, with a wide variety of definitions and contractual features. More basic public-private exchanges involve outsourcing contracts for simple services with limited engagement of public actors in the design and funding of the required activities and investments (Levin and Tadelis 2010). For our purposes, we are interested in forms of *collaboration* that can marshal complementary resources and and engage mutual effort to improve performance (Kivleniece and Quelin 2012).

Crucially, the literature discusses how distinct *types* of private actors can provide distinct types of resources and motivations to create and execute public-private collaborations. *For-profit* organizations, in particular, are said to have higher-powered incentives which can lead to lower costs and higher productivity (Engel, Fischer, and Galetovic 2014). They can also utilize their experience from previous projects, proprietary technology, and extra funding capacity (Fabrizio 2012; Rangan et al. 2006). Yet their higher-powered incentives also create a risk that private operators will pursue excessive appropriation of profits at the expense of service quality and

other relevant externalities, especially in the case of performance dimensions that are more difficult to measure and enforce (Brown and Potoski 2003; Hart, Shleifer, and Vishny 1997).

Another possibility is to partner with *nonprofit* organizations such as institutes, foundations, or civil society organizations (Gazley and Brudney 2007; Valero and Jang 2016; Suarez 2011). Nonprofits contribute with distinct sets of resources and capabilities that are normally not possessed by for-profit organizations. First, they tend to be mission-driven and specialized in particular social issues, leading them to focus on objectives and outcomes that are more aligned with a well-defined group of beneficiaries (McDonald 2007). Second, the incentives of nonprofits are not as high-powered as the incentives of for-profit firms, since their managers do not appropriate a substantial portion of the organization's residual cash flow. These features mitigate the aforementioned hazard of private partners focusing on efficiency at the expense of positive social outcomes (Bennett and Iossa 2009). Thus, compared to collaborations with for-profit firms, public-nonprofit ties tend to increase the perceived *legitimacy* of private engagement (Witesman and Fernandez 2013).

Internal Collaborations between Multiple Public Units

Collaborations, however, do not need to occur only with actors external to the public sector. Far from being monolithic, governments often have a web of specialized structures dealing with diverse activities such as project design, legal enforcement, and service delivery. Especially in the case of governments overseeing large populations, centralized structures have been progressively replaced by more disaggregated structures focusing on narrower, more specialized tasks (Greer 1994)—a process that was reinforced by the NPM emphasis on the infusion of market-like practices into public bureaucracies. Although increased decentralization allowed for more autonomous decision making and localized learning, it also created the challenge of coordinating efforts among diverse public units (Peters 1998; Raab, Mannak, and Cambre 2013).

In this scenario, collaborations across multiple public agencies and units have emerged as the logical approach to plan and execute integrated projects (e.g. Cabral and Krane

Forthcoming). These collaborations can occur either across distinct government units (Agranoff and McGuire 2004), as in the case of partnerships between federal and municipal agencies, or between units of the same government (Sedgwick 2017). Although coordination between these multiple public spheres can be challenging given their distinct objectives and resources, internal collaborations may be seen as more legitimate than external collaborations, since they involve the articulation of public resources and potentially generate spillovers within the state (Andrews and Entwistle 2010; Peters 1998). An emphasis on internal collaborations, however, precludes access to the heterogeneous—and potentially complementary—resources possessed by external organizations. A way to overcome this limitation is to combine internal and external collaborations, a possibility that we examine in our empirical analysis. Another way, discussed next, is to invite and gather input from a broad set of actors involved in the public initiative.

Stakeholder-Orientation

Public units can rely on ideas and suggestions from multiple stakeholders—citizens, civil society associations, entrepreneurs, established firms, and a broad set of contributing actors within the public bureaucracy (Cooper, Bryer, and Meek 2006; Crilly 2011; Buysse and Verbeke 2003). Considering diverse ideas and suggestions can not only increase the odds that public initiatives will result in transformational outcomes, but also help build legitimacy, by attenuating perceptions that these initiatives are simply benefitting public or private actors directly involved in its implementation instead of local communities and other potential beneficiaries (Henisz, Dorobantu, and Nartey 2014; Klein Mahoney, McGahan, and Pitelis. Forthcoming).

From a public governance perspective, this form of engagement also supports a more plural approach to building services and processes that improves key performance dimensions (Osborne 2006). The literature has emphasized, for instance, the role of public leaders in promoting communication channels between public employees and the target population (Denhardt and Campbell 2006). Once inputs are gathered and considered, the whole public bureaucracy must also be receptive to potential adjustments in policy making and implementation, avoiding the usual rigidities in the norms and procedures of the public sector

(Pandey and Scott 2002). Considering all these theoretical elements, we define *stakeholder orientation* as the ability of public units to closely interact with multiple stakeholders even in the absence of formal collaboration, with the overall objective of incorporating novel ideas, adapting policy direction, and adjusting internal processes accordingly.

Based on this initial theoretical framework, our multiple-case methodology, described next, was designed to observe the outcomes of plural forms of interdependent public-private action and to inductively elaborate a configurational theory explaining multiple paths consistent with high performance.

DATA COLLECTION AND METHOD

Case Selection

We built a multiple case sample of 24 public service initiatives in Brazil, India, and South Africa (Table 1). Our selection of countries is appropriate for our research purposes, since they are large emerging economies that suffer from a deficit of essential public services for a substantial portion of their population, combined with many voids related to poor infrastructure and access. At the same time, in these countries we often observe multiple public initiatives to improve public services and infrastructure, generally executed by "pockets" of public bureaucracies conducting socially-oriented innovations (Tendler 1997). In our choice of countries, we started by looking at cases in major emerging economies and then organized dedicated research teams in each country. The presence of these dedicated teams also helped increase their familiarity with the research context, which is considered a crucial step in configurational analysis (e.g. Wagemann & Schneider 2010).

We focused on projects at the municipal level to facilitate the identification of organizational conditions of the public bureaucracy and to allow a more precise, comparative assessment of relative performance across countries. National-, provincial- and state-level jurisdictions vary significantly across countries in terms of the size and responsibilities of public authorities, but conditions and responsibilities at the municipal level tend to be more narrowly defined. This choice also led to the identification of services and activities that are relatively

9

common at the municipal level: education, transportation, urban planning, and bureaucratic services (e.g. services to issue official documents or obtain useful information).

In each sector and country, in line with our previous definition of high performance, we searched for evidence that the project generated improved outcomes in dimensions valued by the target populations (Alford and O'Flynn 2009; Quelin, Kivleniece, and Lazzarini 2017; Moore 1995). As we explain below, we created a common metric to assess all cases. Rather than simply observing whether the initiative was successfully implemented or generated positive results on its own, we also considered evidence of "impact": contrasting outcomes of the considered project *vis-à-vis* comparable groups outside the domain of the project (Kroeger and Weber 2014). Conversely, cases scored poorly when there was either evidence of negative outcomes or when the project was discontinued due to excessive cost or perceived inadequacy (e.g. the initiative was not considered as a viable solution to the target problem). Thus, while the evidence of impact in our successful cases indicates that the initative generated positive outcomes (that is, *high* performance), the cases of failure involve not only the absence of such positive results but also poor outcomes adversely affecting the target population (i.e. *low* performance). To find potential cases, we conducted a broad review of evaluation studies of public service initiatives in existing repositories (such as the World Bank impact evaluation library) and contacted contacted sector specialists to explore suggestions for other potential cases with evidence of positive or negative outcomes. For more detailed information of our cases and their evidence of impact, see Table A1 of the supplementary appendix.

<center>*<Table 1 around here>*</center>

Method

Starting with those 24 cases and our baseline theoretical framework, we sought to unveil novel explanatory relationships emanating from the comparative analysis of high and low performance cases. Specifically, we adopted a *configurational comparative method* (Dwivedi, Joshi, and Misangyi 2018; Fiss 2011; Misangyi and Acharya 2014; Ragin 2006; Thiem and Baumgartner 2016; Thiem and Dusa 2013) to examine organizational patterns consistent with

high performance. This is possible by systematically comparing different cases with a set of defining characteristics or *conditions*, which may be associated with an outcome of interest (Rihoux and Ragin 2009). An important feature of configurational methods is that they account for *equifinality*, i.e. they accommodate situations where multiple paths can lead to the same outcome (Fiss 2007; Ragin 2008; Rihoux and Ragin 2009). Thus, equifinality is consistent with our multiple-actor, multiple-logic perspective to explain high public performance, as discussed in the previous theory section.

To come up with distinct configurations, we apply *fuzzy-set qualitative comparative analysis* (fsQCA). This method employs fuzzy-set algebra to find combinations of conditions consistent with each outcome. In fsQCA, instead of coding the simple presence or absence of a condition, it is possible to consider the *degree* to which a condition is present (e.g. the *intensity* of collaboration or the extent of public operational capacity). Thus, fsQCA allows the researcher to examine not only conditions that will explain the occurrence of a phenomenon—i.e., if a condition is present (referred to as "fully in") or absent ("fully out")—but also the extent to which each observation is consistent with high or low performance. For 24 cases, Marx and Dusa (2011) recommend no more than five conditions, which is indeed what we unveiled in our previous theoretical discussion: public capacity, the three collaboration types, and stakeholder orientation.

In light of the analysis of the cases, our goal was to "reevaluate theoretical domains in a configurational manner" (Misangyi, Greckhamer, Furnari, Fiss, Crilly, and Aguilera 2017, 268), a research process referred to as *theory elaboration*. Yet, because this theory elaboration effort is anchored to our limited ("small-N") set of cases, our results are not generalizable (Greckhamer, Misangyi, and Fiss 2013). Our novel theorizing is therefore *midrange*, that is, based on an effort to unveil complex configurations of theoretical conditions applied to our specific empirical context (for further discussions and other applications, see, for instance, Andrews and Entwistle 2010; Crilly 2011; Raab, Mannak, and Cambre 2013; Sedgwick 2017; Thiem and Baumgartner 2016; Thiem and Dusa 2013; Thomann, van Engen, and Tummers 2018; Wang 2015).

Data Collection

We collected data through qualitative interviews anchored on customized scales (rubrics) to measure our conditions (see Table A2 of the supplementary appendix). Our use of rubrics instead of agreement (Likert) scales is justified because we wanted to guarantee comparability across all responses. In our rubrics, we carefully described what each point in the scale meant for each item, thus anchoring responses on distinct types of behavior corresponding to different levels of the construct (Oakleaf 2009). For instance, the highest score of our performance measure (5) involves a situation of strong quantitative evidence that the target populations improved on key outcome variables beyond what was observed in comparable groups. The intermediate score (3), in turn, would reflect *some* presence of the condition; that is, we can still consider that the condition is present, but in weaker form. In our outcome rubric, this is expressed as some evidence of positive outcomes, even though results are not totally aligned (e.g. some performance indicators improve while others have a more ambiguous development). Finally, the lowest score of our rubric (1) would expose a situation of clear low performance— e.g. evidence that project failed due to strong opposition and/or evidence of weak or even negative outcomes.

Dedicated research teams in each country, with a deep understanding of the local context, performed the interviews and overall data collection. For each project, they conducted three interviews with public officials in strategic and/or operational positions, and with managers of private organizations or general stakeholders that conceived, implemented, or studied the project (see Table A3 of the supplementary appendix). The selection of multiple interviewees allowed us to mitigate common respondent bias and check the reliability of their assessments. Each interview lasted at least one hour; in total, our interviewing process took about 72 hours, not including the field trips and additional interviews required to validate all the cases. Besides collecting coded information following our rubrics, we also gathered data from official reports, articles, and field trips when possible, paying attention to potential differences in the interpretation of the phenomena under analysis.

Thresholds and Calibration

In configurational analysis, researchers have to code whether each condition (say, public capacity or collaboration between public units) and the overall outcome of interest (high performance) are present ("in") or absent ("out") in a particular case (Misangyi et al. 2017). This process is referred to as *calibration*. A natural way to calibrate our measures would be to directly examine responses to each rubric. Given that responses to our rubrics range between 1 and 5, we could for instance suppose that a score of 3 would represent moderate presence ("more in than out") and so on, successively. However, recall that, to increase the reliability of our measures, we used multiple raters and, as we explain below, some of our measures involve multiple items to capture distinct complementary dimensions.

We thus created our *raw data matrix* based on composite measures as the average scores of the raters for multiple items (see Table A4 in the supplementary appendix). For example, in the Sobral Education Program (Brazil), the condition "collaboration with for-profit organizations" had a final composite score of 2.84. The raw measure of this condition is the average of the scores of two items coding interdependent effort (to be explained below): the first item received scores of 4, 3, and 2 from the three interviewees (3 on average), while the second item, scores of 3, 3, and 2 (2.67 on average). With these raw database (composite measures), we then applied the *recoding calibration method* (Emmenegger, Schraff, and Walter 2014; Schneider and Wagemann 2012; Thomann et al. 2018). Roughly speaking, the method allows us to consider the raw scores coming from the survey instrument (in our case, interviews using rubrics) and then define the extent to which a particular condition is present or not in each case. Tables A6 and A7 of the supplementary appendix provide more details on our final calibration and its procedures.

Measuring the Performance of the Public Initiative

We compiled information on project-level outcomes and wrote two-page reports for each case summarizing all relevant performance-related information (including the conclusions of studies assessing the impact of each project, as discussed before). Based on these reports, we then asked three independent raters to provide their suggested performance scores on a 1-5 scale. Differently from the coding of conditions, we opted to use independent raters instead of assessments from project participants in order to anchor our analysis of performance based on impact studies and other independent sources of data. Agreement across raters was very high (Cronbach's Alpha and ICC equal to 0.949, and Cohen's Kappa agreement test with $p < 0.001$). The average of raters' responses was then used as our final outcome measure of performance (evidence of positive outcomes likely valued by the target populations). The average score of our high performance cases ("more in than out" or "fully in") was 4.5, whereas the average score of the low performance cases was 1.44.

Measuring the Theoretical Conditions

Differently from the measurement of the performance outcome, all conditions were gauged based on each rubric and the input of three actors who were highly knowledgeable of each project (e.g. local managers or public officials). Their familiarity with project-level factors increased our confidence that our theoretical conditions were properly assessed (Wagemann & Schneider 2010). After completing the set of three interviews for each of the 24 cases, we performed reliability tests to ensure consistency within cases as well as agreement across interviewees for the same case (see Table A5 of the supplementary appendix).

Public Operational Capacity. Following our theory discussion, we measured public operational capacity as a five-item composite measure including i) public leaders' articulation of desired goals, ii) merit-based (as opposed to politically-motivated) staffing, iii) extent of monitoring and accountability, iv) existence of processes to facilitate implementation and change, and v) adoption of anti-corruption practices (Cronbach's Alpha = 0.93). In this case, the highest scores in the rubrics (5) generally indicate a situation where leaders have clear goals, hire skilled managers, and promote strong monitoring and accountability mechanisms (Andrews et al.

2017; Barzelay and Armajani 1992; Miller and Whitford 2016). The intermediate points (3) reflect some presence of public capacity, even though some features remain underdeveloped—for instance, although the leader has a clear vision of the project, goals are not clearly articulated and expressed to the team. Finally, the lowest scores (1) indicate poor leadership, use of political appointments, and a lack of accountability.

Collaboration. In line with our previous theory discussion, we focused on three main forms of collaboration, namely: between public units, with for-profit organizations, and with nonprofit organizations. For each type of partnership, we relied on two questions: We coded whether the public agency responsible for the project i) mutually collaborated with the partner, and ii) recruited/engaged people from the partner to work on the project. Consistent with our theory discussion on forms of collaboration, these questions were particularly designed to gauge the *interdependence* of the collaborating parties (e.g. Mahoney et al. 2009; Sedgwick 2017).

For each collaboration type, we then calculated a composite measure based on the average of those two items. We thus arrived at three composite measures for each collaboration type: between public units, with for-profit organizations, and with nonprofit organizations. Given that Cronbach's Alphas are not usually recommended for two-item measures, we checked their reliability using Spearman-Brown correlations (Eisinga, Te Grotenhuis, and Pelzer 2013); they were respectively 0.73, 0.70 and 0.85, all above or equal to the acceptable level of 0.70.

Stakeholder Orientation. We measured the stakeholder orientation of the public unit with four items, coding i) the permeability of leaders to external suggestions from citizens and other stakeholders, ii) the ability of leaders to develop external channels of communication with stakeholders, and iii) the existence of rigid formal rules constraining the adoption of external ideas (reverse scored), iv) the existence of internal norms and procedures facilitating improvement and change (Cronbach's Alpha = 0.73). In this condition, in line with our theory discussion (Denhardt and Campbell 2006; Pandey and Scott 2002; Raab et al. 2013), the highest scores (5) express a condition where top leadership communicates with external stakeholders and the internal processes of the public bureaucracy are highly permeable to inputs from these

external actors. The intermediate points (3), in turn, reflect some extent of stakeholder orientation, such as when the leader centralizes decision making but there is some degree of interaction with managers, external communication, and adaptation to new suggestions. The lowest scores (1), in contrast, characterize total absence of stakeholder orientation: leaders rarely involve multiple stakeholders in the decision-making process and organizational inflexibilities hamper any possible change and improvement.

Analytical Procedures

Our primary objective is to analyze whether the presence or absence of any theoretical condition (alone or combined) is sufficient to observe superior performance. Thus, after calibration, we defined the consistency and frequency thresholds for the fsQCA. Consistency refers to the degree to which a particular case that exhibits a particular configuration is also associated with the final outcome of interest (Ragin 2006). In our case, consistency thus codes the percentage of cases with a certain configuration that have evidence of high performance. We conservatively adopted a 0.800 consistency threshold (Ragin 2006). This threshold was defined after examining the *truth table* (see Table A8 of the supplementary appendix), which indicates distinct configurations and the high-performance cases that are consistent with these configurations. Also, given our relatively small sample size, we followed the recommendation by Rihoux and Ragin (2009: 107) and considered configurations with at least one representative case (this is the so-called *frequency threshold*). That is, we focused on logical configurations that were observed in our sample of high-performance cases. Finally, in our analysis, we considered the *intermediate solution* generated by the fsQCA software, that is, we considered all possible logical combinations independently if each combination is in our sample or not (Fiss 2011; Misangyi and Acharya 2014).

<div align="center">

RESULTS: CONFIGURATIONS AND CASE ILLUSTRATIONS

</div>

Main Results

Table 2 presents the results of our sufficiency analysis, showing three solutions consistent with high performance (1a, 1b and 1c). As is usual in the QCA literature (Fiss 2011; Misangyi

and Acharya 2014), we interpret the configurations as follows: central conditions are represented by "●" (presence) and "⊗" (absence); while contributing conditions are marked with "●" (presence) and "⊗" (absence). Moreover, following Greckhamer (2016), *necessary* central conditions are represented by "■" (presence), while *necessary* contributing conditions are represented by "■" (presence). These necessary conditions come from the *necessity analysis* reported in Table A10 of the supplementary appendix, following Dwivedi et al. (2018) and Ragin (2000, 2008). In a nutshell, necessary conditions appear in all configurations consistent with a given outcome and thus can be considered as a pre-requisite for either high or low performance. Blank spaces indicate a "don't care" situation, that is, the condition is not relevant to that configuration. Central conditions are present in both parsimonious and intermediate solutions, whereas contributing conditions are only present in the intermediate solution (Fiss 2011: 403).[1] In each configuration, we also select *exemplary cases* via an in-depth examination of our interviews and detailed descriptions of the public initiatives in the sample (Byrne and Ragin 2009: 231).

<center><Table 2 around here></center>

Looking at the set of solutions (configurations 1a, 1b, 1c), and consistent with our plural approach, we find two general paths to a high-performance public initiative. The first path involves higher *internal* engagement (configuration 1a), as it is centered on interactions between actors within the public sector, while the second consists of higher *external* engagement (with two configurations, 1b and 1c), thus with higher externalization effort. Sectors and countries are fairly represented in all configurations (see Tables A12 and A13 of the supplementary appendix). Below, we describe our revealed solutions in more detail, triangulating between the results of our configurational analysis, illustrative quotes from our interviews and other sources of information,

---

[1] Technically speaking, central conditions are present if we consider the existing possible configurations within the sample (*parsimonious* solution), or above and beyond the sample (*intermediate* solution). While, contributing conditions are present only in the possible configurations within the sample (*parsimonious* solution) (Fiss 2007; Rihoux and Ragin 2009).

as a way to improve our understanding of how conditions work and complement one another in each configuration.

Path with higher internal engagement. Configuration 1a of Table 2 indicates that, in the absence of collaborations with nonprofit organizations, high-performance public initiatives can work through a combination of partnerships within the public sector (collaborations between distinct public units), public operational capacity, and stakeholder orientation. We label this path as *internal* because it mostly involves actions inside the public sector, broadly defined (even though the configuration also displays collaborations between multiple public units).

To increase our understanding of this result and gain more insights on this configuration and its mechanisms, we selected an exemplary case, India's New Delhi Metro, to illustrate our finding. New Delhi Metro was the second metro rail project in India. Significant social and economic outcomes resulted from the project—for example, a reduction in the total number of vehicles on the streets and savings due to efficient implementation. As seen in the following quote, there is evidence of strong public capacity, defined as a combination of effective leadership and execution:

> *Leadership played a very important role in the success of the Delhi Metro… In the case of Delhi metro, it played a very important role especially in terms of getting it through the bureaucracy, getting the right talent, starting implementation and taking faster decisions. For instance, the Delhi Metro was able to pull out a lot of good people from Railways who were extremely qualified in terms of their subject domain… (Expert in public transportation, pers. comm.)*

The project also relied on collaborative action within the public sector. Indeed, it even resulted in the creation of a new public organization resulting from the joint effort of multiple government units, The Delhi Metro Rail Corporation (DMRC). DMRC was co-owned by the Government of India and the Government of National Capital Territory of Delhi. Delhi Metro was also the first project designated to the Ministry of Urban Development, beyond the core activities of the state-owned firm Indian Railways. Acting as a focal unit for the project, DMRC helped harmonize the goals of those distinct public units and create internal processes for project

execution. In addition, during the planning phase of the project, regular stakeholder consultations were held; these consultations were critical to assess service attributes that final users would demand, as well as other important considerations that DMRC would have to incorporate in the project. This feature reinforces the presence and importance of stakeholder orientation. Although some collaborations with for-profit private firms did occur, the project was fundamentally built through concerted efforts within the public sector, reinforced by strong leadership with openness to receive input from multiple actors.

Path with higher external engagement. Configurations 1b and 1c are *neutral permutations* (Fiss 2011), i.e. they share the same central conditions but differ in the contributing ones. We interpret both configurations as a similar path—considering that "the permutations do not affect the overall performance of the configuration" (Fiss 2011:398)—while at the same time explaining their observed variations. In this sense, configurations 1b and 1c can be thought of as multi-collaborative configurations involving both for-profit and nonprofit organizations, combined with the necessary presence of strong public operational capacity. However, while in configuration 1b stakeholder orientation is a contributing condition, in configuration 1c stakeholder orientation is a "don't care" condition. The exactly opposite is true with respect to collaboration between public units: in configuration 1b, it is a "don't care" condition while it is a contributing condition in configuration 1c..

An exemplary case for configuration 1b is South Africa's Siyakha Nentsha project (translated as "building with young people"), an education program for teenagers in South Africa's KwaZulu-Natal province. The project involved developing extracurricular skills among young students, with the objective of teaching them how to mitigate threats from HIV/AIDS—whose incidence in this region of South Africa is particularly high. There is evidence that the project resulted in significant improvements in safe sexual behavior, knowledge of HIV/AIDS risks, and diverse cognitive abilities. The project was implemented via important collaboration with nonprofit organizations, chiefly Population Council, a body specialized in conducting research and delivering public health solutions, particularly in emerging economies. Population

Council nurtured another nonprofit organization, as an independent, local, and mission-driven agency, called Isihlangu Health and Development Agency. The nonprofit organizations worked with the government, more specifically, the Department of Education, to offer specialized education programs for HIV prevention. The collaboration with formal government systems allowed the program's participants to include their training in the public education records, improving their opportunities in future job searches.

Collaboration with for-profit organizations, in turn, came through the activities of AccuData, a for-profit firm specialized in data and research solutions with expertise in South Africa. These two types of collaborations apparently complemented each other in terms of the types of resources and capabilities that they brought, with AccuData providing technical capabilities and Population Council (as well as other nonprofits) providing specialized knowledge of health and educational programs, as well as capabilities to increase the engagement of local communities. Nonprofits were also key actors in selecting for-profit partners, in a way that was considered aligned with the social objectives of the project. Besides these collaborations, nonprofit managers also organized meetings with citizens of various ages, with the objective of learning their views about the project and the vulnerabilities of the target group. The project also recruited youths from local municipalities, who received basic salaries and training, apparently making their parents more predisposed to support the initiative. Thus, reflecting the presence of stakeholder orientation, the project had not only to marshal community resources but also be sensitive to their inputs:

> *Because we were in a tribal area, we had to go to different routes... They made a rule that we had to hire people from the community... We had to be respectful and go to them, be clear and negotiate (Researcher specialized in education, pers. comm.).*

While this example shows the importance of stakeholder orientation in configuration (1b), the other configuration with higher external engagement (1c) adds the presence of collaboration between public units as a condition leading to high performance. Indeed, in this configuration we see a very plural combination of *all* types of internal and external collaboration

that we reviewed in our theory section. An exemplary case is the bus rapid transport (BRT) system in the city of Curitiba, Brazil, designed to provide fast, low-cost transportation to local citizens, later adopted in several countries (Lindau, Hidalgo, and Facchini 2010). Nonprofit education and research institutes helped with studies and provided technical specialists to work on the project, while for-profit manufacturers of vehicles such as Sweden's Volvo contributed with distinct technical knowledge to design customized buses and related transport technologies. The project also relied on intense collaboration between public units involving multiple state actors and state-owned specialized organizations. A leading actor involved in the project was IPUCC (Institute of Urban Research and Planning of Curitiba), a public unit responsible for the planning and monitoring of metropolitan activities, based on general directives set by the municipal government and in collaboration with the local transport authority and other government spheres:

> *… the team of IPUCC who designed all the projects, all final projects, visual communication, urban infrastructure, how the system would work… [Also] Brasilia [the capital city of the federal government] helped us a lot in this process, you know, the people who were in charge of financing these projects were always key partners (Former Secretary of Planning and Mayor, pers. comm.).*

There is also evidence that these collaborations between public units effectively functioned as conduits of new ideas on how to improve service delivery. IPUCC held a multifunctional team involving engineers, sociologists, and urban planners closely collaborating with the transport authority of the municipal government on a broad range of activities. Furthermore, IPUCC closely collaborated with URBS, a government agency specialized in urban transport. In other words, widespread public collaborations in Curitiba's BRT System helped incorporate and articulate relevant contributions from myriad public units, essentially functioning as a mechanism to foster permeability to the input of multiple (internal) stakeholders.

Configurations Associated with Low Performance

One of the premises of QCA analysis is that configurations leading to high performance are not necessarily the symmetric opposite of configurations leading to low performance. Thus, it

is also informative to examine combinations consistent with low performance. As seen in Table 3, three out of four (2a, 2b, and 2d) low-performance solutions involve the absence of public operational capacity, which is aligned with our previous finding that this attribute represents a key driver of high performance. Yet configuration 2c shows that even if public operational capacity is present and combined with stakeholder orientation, the absence of *all* collaboration types leads the initiative to fail. In other words, although public operational capacity seems to be a necessary condition to high performance, its presence does not necessarily avoid a low performance result.[2] However, solution 2d indicates that the absence of public operational capacity is associated with low performance even when all other conditions are present. Finally, configuration 2a, with the greatest number of representative cases (5), shows that the combined absence of multiple conditions is a major driver for low performance. In sum, the results of configurations associated with low performance not only reinforce the centrality of public operational capacity, but also confirm the potential role of plural collaborations in fostering the performance of public initiatives.

<<*Table 3 around here*>>

DISCUSSION: UNDERSTANDING PLURAL FORMS OF INTERDEPENDENT PUBLIC-PRIVATE ACTION

In light of our empirical findings, in this section we elaborate a midrange theory of how public capacity and plural forms of collaboration create positive outcomes to target populations. Figure 1 depicts the conditions associated with our unveiled paths—that is, combinations of complementary conditions leading to high performance—whose underlying interactions are explained below and summarized in Table A15 of the supplmenetary appendix.

<*Figure 1 around here*>

---

[2] This interpretation is further supported by a *necessity analysis of the failure* (Dwivedi et al, 2018; Ragin, 2000, 2008) shown in Table A11 of the supplementary appendix (the corresponding truth table is in Table A14). As Table A11 shows, the joint absence of all collaborations is usually associated with a low performance at a 5% significance level.

In line with the equifinality principle in configurational analysis (Fiss 2007; Ragin 2008), we unveil two general paths consistent with superior performance. The first path (corresponding to Table 2, solution 1a) relies heavily on internal public effort: collaborations between the focal unit and other public units, as well as a general predisposition to incorporate inputs from external stakeholders, in a context where public-nonprofit collaborations are scarce. In this configuration, the focal public unit fosters collaborations between multiple government units, thus increasing the ability to seamlessly articulate distinct public capabilities, coordinate all necessary changes, and improve delivery (Andrews and Entwistle 2010; Cabral and Krane Forthcoming). Also, leaders encourage public managers to adopt mechanisms to dialogue with stakeholders and incorporate their input (Crilly 2011; Buysse and Verbeke 2003). Yet simply being open to external suggestions can be innocuous if not accompanied by organizational efforts to set goals as a function of those new suggestions and to adapt internal processes accordingly.

In addition, the central ability to connect with and incorporate suggestions from stakeholders reduces the need for formal collaboration with private actors, especially nonprofit organizations. In our previous theory discussion, we argued that an advantage of involving nonprofits is that they are often specialized in certain types of beneficiaries and critical social issues. Therefore, high public capacity combined with strong stakeholder orientation amplifies the predisposition of leading public actors to listen to and understand the needs of target populations as well as their ability to adapt to the specificities of the social activity. The use of internal resources and a general predisposition to receive novel input also increases perceptions of legitimate action (Cooper, Bryer, and Meek 2006; Klein et al. Forthcoming; Raab et al. 2013). Consequently, with these conditions in place, superior performance can naturally flow from actions that occur mostly inside the state. Notice, however, that this configuration does not involve total absence of collaboration. In this path, alliances between public units *within* the public bureaucracy promote joint action across distinct public units with complementary roles in the process of service delivery or sharing a common interface with beneficiaries.

Our second path, in contrast, involves a higher emphasis on private actors, both for-profit and nonprofit (Table 2, solutions 1b and 1c). This path thus expresses a multiple-actor integrative effort targeting community-level outcomes (Provan and Milward 2001). In our theory review, we argued that, despite their capital and knowledge-based benefits, the engagement of for-profit firms may pose a risk to social outcomes if they place excessive emphasis on the appropriation of economic gains (Cabral et al. 2013). A plural collaborative approach involving the joint and central presence of nonprofit organizations will tend to attenuate this risk (Gazley and Brudney 2007). Mission-driven nonprofits, in this context, help guarantee that the project will not substantially drift from the original (social) objectives of the public initiative. In addition, nonprofits often have a comparative advantage in important dimensions of public service, such as specialized knowledge of target beneficiaries, skills to communicate with local communities, and even extra philanthropic resources to support interventions. In other words, in the path involving external engagement, nonprofit organizations tend to have a key role, helping bridge the focal public unit and the external partners, contributing with specialized knowledge and increasing the legitimacy of externalization.

Although the presence of external collaborations allows for the incorporation of novel and valuable input, the focal public unit will still need to consider and articulate multiple contributions to the project. In situations involving moderately plural collaborations with for-profit and nonprofit organizations—i.e., without the required presence of collaborations with other public units—stakeholder orientation remains an important condition to access novel input, including potential contributions coming from other relevant units of the public bureaucracy. In this case, given the emphasis on external collaborations associated with the presence of stakeholder orientation, collaborative effort between the focal unit and other public units is not required (i.e. it is a "don't care" condition).

In contrast, situations involving highly plural configurations that include both external *and* internal collaborations allow for a broad and diverse array of inputs to successfully craft and adjust performance-enhancing processes, thereby making the presence of stakeholder orientation

less necessary (i.e. now the "don't care" condition is stakeholder orientation). In this case, collaborations between public units can functionally act as a mechanism of stakeholder orientation when multiple public units contribute with distinct ideas and resources to improve project design and adaptation. These mechanisms explain the neutral permutation (Fiss 2011) between internal collaborations and stakeholder orientation, discussed in the previous section. However, even when required as a condition in the external engagement path, stakeholder orientation is not as central as in the internal engagement path, where public actors must manage relationships across different public service units.

In both paths, as long suggested by public management scholars (e.g. Moore 1995; Miller and Whitford 2016), public operational capacity plays a key role. In other words, externalization is far from a process where private actors supplant inefficient or incapable governments (e.g. Deaton 2013). Rather, public capacity is a central condition to complement and improve the performance-enhancing effect of externalization (Alford 2015). Moreover, regardless of whether public initiatives emphasize internal or external effort, they first need the full engagement of political leaders who set a given collaborative agenda and define a common vision guiding the mobilization of public and private resources. Mechanisms to promote monitoring and accountability will also avoid the risk that new projects will fall prey to corruption and misallocation of resources. Thus, although our plural framework builds on a multiple-actor, multiple-logic perspective (Bryson et al. 2017), the role of processes and capabilities *within* the public bureaucracy remains central to explain superior performance.

CONCLUDING REMARKS

Contributions to Debate on Plurality and Public Performance

Our empirical results and our theory elaboration effort contributes to the literature in several important ways. Connecting with the PVM perspective, we examine conditions that lead a public initiative to achieve high performance, defined as positive outcomes that are valued by its target population (e.g. Alford and O'Flynn 2009; Moore 1995). Crucially, we adopt a plural approach (Lynn Jr et al. 2000; Osborne 2006) by proposing how multiple actors—public, for-

profit, and nonprofit—can engage in multiple logics of action leading to high performance (Bryson et al. 2017). Our configurational analysis (Fiss 2011) provides a structured procedure to unveil these multiple logics of action combining multiple actors, under the premise that high performance may be achieved via distinct organizational paths (i.e. equifinality). By describing distinct paths consistent with superior performance, with higher public or higher private engagement, we thus demonstrate how public initiatives can be successfully organized in diverse ways and with distinct logics of action—a topic that has received attention in organizational research (e.g. Misangyi 2016). Therefore, instead of examining or proposing single organizational solutions to complex public problems—such as more externalization or more governmental action—our work suggests that policy makers should consider alternative paths to engage public and private actors in distinct and complementary ways.

Importantly, although externalization is usually seen as way to benefit from the input of actors outside the public bureaucracy, we argue that public actors can develop external channels of interaction even if they do not nurture *formal* external collaborations. We theorized and found that the presence of stakeholder orientation is especially relevant when the public initiative emphasizes internal public action. Permeability to ideas coming from a host of diverse actors becomes crucial not only to improve service design but also to more holistically engage multiple beneficiaries (Henisz et al. 2014; Klein et al. Forthcoming). Furthermore, although some scholars have proposed that the engagement of private actors can be a response to poor public capacity (Auerswald 2009; Krasner and Risse 2014), we contend that strong public capacity is critical to extract greater value from private sector contributions during collaborative projects. Our two proposed paths require the presence of capable public leaders who clearly express their goals, engage capable managers, and closely follow the outcomes of the initiative (Andrews et al. 2017; Barzelay and Armajani 1992; Miller and Whitford 2016). Thus, our results clearly show that internal public capacity *complements* externalization. Private action in public projects, even when desirable, is no substitute for weak governments.

Limitations, Generalizability of our Results, and New Directions

There are several ways in which future research could build upon our findings and improve the understanding of how organizational configurations influence the performance of public initiatives. An important issue is if there are other conditions that may affect project outcomes, besides our theoretically informed factors. Indeed, after assessing our configurational solutions, we reexamined our cases to identify other conditions that might explain high performance. We only found additional factors that were highly idiosyncratic (project-specific), instead of general conditions that could influence the success of multiple projects (see Table A16 in the Supplementary Appendix). In-depth, qualitative studies may help scrutinize other potential conditions and their underlying dynamics in high performance projects.

In addition, we have no intention to claim that our results are generalizable, as the configurational analysis literature emphasizes that the key objective of the method is to find logical combinations of conditions. Thus, Rihoux and Ragin (2009: 12) argue that "generalization in QCA studies is best conceptualized as 'modest'." In this sense, our theory elaboration effort tries to propose theoretically consistent configurations emanating from our examined projects. Therefore, we would welcome future research to assess and even possibly test the relevance of our proposed conditions using a larger number of cases in multiple sectors, countries, and government levels. In particular, regression analysis can complement our QCA results by examining not only the joint occurrence of our proposed conditions, but also whether ther interactions statistically explain superior performance.

Given that we focus on large emerging economies, a related concern is whether our paths can be found in distinct institutional contexts, such as in the case of developed countries. Arguably, there might be less variation in terms of public capacity in more developed economies, whose improved institutions may increase public sector accountability and execution capabilities (Andrews et al. 2017; Deaton 2013). Yet, even in those countries, effective implementation may not only require public capacity but also a host of distinct and diverse organizational mechanisms to interact with external actors. For instance, Villani, Greco and Phillips (2017) describe two cases of public-private collaboration in the Italian health sector, whose diverging outcomes can

be traced back to the capacity (or lack thereof) to implement knowledge-sharing and interdependent processes at the public-private boundary. Thus, we believe that our proposed paths represent logical combinations of conditions that are plausible and relevant in developed and emerging economies alike—which is not to say that they represent an exhaustive set of solutions. Novel plural configurations may emerge from cases that incorporate a more diverse set of contextual conditions, potentially improving our understanding of how public initiatives work and generate positive outcomes that are highly valued by their target populations.

REFERENCES

Agranoff, Robert, and Michael McGuire. 2004. *Collaborative Public Management: New Strategies for Local Governments*. Washington: Georgetown University Press.

Alexander, Jennifer, and Renée Nank. 2009. Public-Nonprofit Partnership Realizing the New Public Service. *Administration & Society*, 41(3): 364-386.

Alford, John, and Janine O'Flynn. 2009. Making sense of public value: Concepts, critiques and emergent meanings. Intl Journal of Public Administration, 32(3-4): 171-191.

Alford, John, and Janine O'Flynn. 2012. *Rethinking Public Service Delivery: Managing With External Providers*. London: Palgrave Macmillan.

Alford, John. 2015. Weighing the Public Value of Alternative Methods of Providing Public Services: Toward a Contingency Framework. In *Public Value and Public Administration*, Bryson JM, Crosby BC, Bloomberg L (eds). Washington: Georgetown University Press, p. 68-81.

Andrews, Matt, Lant Pritchett, and Michael Woolcook. 2017. *Building State Capability: Evidence, Analysis, Action.* Oxford: Oxford University Press.

Andrews, Rhys and Tom Entwistle. 2010. "Does Cross-Sectora Partnership Deliver? An Empirical Exploration of Public Service Effectiveness, Efficiency, and Equity." *Journal of Public Administration Research and Theory* 20(3):679–701.

Andrews, Rhys, Malcolm James Beynon, and Aoife Mary McDermott. 2015. Organization Capability in the Public Sector: A Configurational Approach. *Journal of Public Administration Research and Theory*, 26(2): 239–258.

Auerswal, Philip. 2009. Creating Social Value. *Stanford Social Innovation Review*, 7(2): 50-55.

Barzelay, Michael, and Babak J Armajani. 1992. *Breaking Through Bureaucracy: A New Vision for Managing in Government*. Berkeley: University of California Press.

Bell, R. Greg, Igor Filatotchev, and Ruth V. Aguilera. 2014. Corporate Governance and Investors' Perceptions of Foreign IPO Value: An Institutional Perspective. *Academy of Management Journal*, 57(1): 301–320.

Bennett, John, and Elisabetta Iossa. 2009. Contracting Out Public Service Provision to Not-For-Profit Firms. *Oxford Economic Papers*, 62: 784-802.

Brown, Kerry A., and Stephen P. Osborne. 2012. *Managing Change and Innovation in Public Service Organizations*. Abingson: Routledge.

Brown, Trevor L., and Matthew Potoski. 2003. Transaction Costs and Institutional Explanations for Government Service Production Decisions. *Journal of Public Administration Research and Theory*, 13(4): 441-468.

Bryson, John M., Barbara C. Crosby, and Laura Bloomberg. 2015. Discerning and Assessing Public Value: Major Issues and New Directions. In *Public Value and Public Administration*, Bryson JM, Crosby BC, Bloomberg L (eds). Washington: Georgetown University Press, p. 1-21.

Bryson, JohnM., Alessandro Sancino, John Benington, Eva Sørensen. 2017. Towards a multi-actor theory of public value co-creation. *Public Management Review*, 19(5): 640–654.

Buysse, Kristel, and Alain Verbeke. 2003. Proactive Environmental Strategies: A Stakeholder Management Perspective. *Strategic Management Journal*, 24(5): 453-470.

Byrne, David and Charles C. Ragin. 2009. *The Sage Handbook of Case-Based Methods*. London, Los Angeles, New Delih, and Singapore: Sage Publications.

Cabral, Sandro, and Dale Krane. Forthcoming. Civic Festivals and Collaborative Governance. *International Review of Administrative Sciences*.

Cabral, Sandro, Sérgio G. Lazzarini, and Paulo Furquim de Azevedo. 2013. Private Entrepreneurs in Public Services: A Longitudinal Study of Outsourcing and Statization in Prisons. *Strategic Entrepreneurship Journal*, 7(1): 6-25.

Campbell, Joanna Tochman, David G. Sirmon, and Mario Schijven. 2016. Fuzzy Logic and the Market: A configurational approach to investor perceptions of acquisition announcements. *Academy of Management Journal*, 59(1): 163–187.

Collier, Paul, and David Dollar. 2002. Aid Allocation and Poverty Reduction. *European Economic Review*, 46(8): 1475–1500.

Considine, Mark, Jenny M. Lewis, and Damon Alexander. 2009. *Networks, Innovation and Public Policy: Politicians, Bureaucrats and the Pathways to Change Inside Government*. Houndmills: Palgrave McMillan.

Cooper, Terry L., Thomas A. Bryer, and Jack Wayne Meek. 2006. Citizen-Centered Collaborative Public Management. *Public Administration Review*, 66(s1): 76-88.

Crilly, Donal. 2011. Predicting Stakeholder Orientation in the Multinational Enterprise: A Mid-Range Theory. *Journal of International Business Studies*, 42(5): 694–717.

Deaton, Angus. 2013. *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton University Press,.

Denhardt, Janet V., and Kelly B. Campbell. 2006. The Role of Democratic Values in Transformational Leadership. *Administration & Society*, 38(5): 556-572.

Dwivedi, Priyanka, Aparna Joshi, and Vilmos F. Misangyi. 2018. Gender-Inclusive Gatekeeping: How (Mostly Male) Predecessors Influence the Success of Female Ceos. *Academy of Management Journal*, 61(2): 379–404.

Eisinga, Rob, Manfred te Grotenhuis, and Ben Pelzer. 2013 The Reliability of a Two-Item Scale: Pearson, Cronbach, or Spearman-Brown? *International Journal Of Public Health*, 58(4): 637-642.

Emmenegger, Patrick, Dominik Schraff, and André Walter. 2014. *QCA, the Truth Table Analysuis and Large-N Survey Data: The Benefits of Calibration and the Importance of Robutness Test*. Compass Working Paper Series.

Engel, Eduardo, Ronald D. Fischer, Alexander Galetovic. 2014. *The Economics of Public-Private Partnerships: A Basic Guide*. Cambridge: Cambridge University Press.

Fabrizio, Kira R. 2012. Institutions, Capabilities, and Contracts: Make or Buy in the Electric Utility Industry. *Organization Science*, 23(5): 1264-1281.

Fiss, Peer C. 2007. A Set-Theoretic Approach To Organizational Configurations. *Academy of Management Review*, 32(4): 1180–1198.

Fiss, Peer C. 2011. Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research. *Academy of Management Journal,* 54(2): 393–420.

Fredriksson, Anders. 2014. Bureaucracy intermediaries, corruption and red tape. *Journal of Development Economics*, 108(May): 256-273.

Fukuyama, Francis. 2004. *State-Building: Governance and World Order in the 21st Century*. Ithaca: Cornell University Press

Gazley, Beth, and Jeffrey L. Brudney. 2007. The Purpose (and Perils) of Government-Nonprofit Partnership. *Nonprofit and Voluntary Sector Quarterly*, 36(3): 389-415.

Geddes, Barbara. 1994. *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley: University of California Press.

Greckhamer, Thomas, Vilmos F. Misangyi, and Peer C. Fiss. 2013. The Two QCAs: From a Small-N to a Large-N Set-Theoretic Approach. in PC. Fiss, B Cambré, A Marx (ed.). *Configurational Theory and Methods in Organizational Research (Research in the Sociology of Organizations, Volume 38).* Emerald Group Publishing Limited.

Greckhamer, Thomas. 2016. CEO Compensation in Relation to Worker Compensation Across Countries: The Configurational Impact of Country-Level Institutions. *Strategic Management Journal,* 37, 793–815.

Greer, Patricia. 1994. *Transforming Central Government: The Next Steps Initiative*. Buckingham: Open University Press.

Hart, Oliver D., Andrei Shleifer, and Robet W. Vishny. 1997. "The Proper Scope of Government: Theory and an Application To Prisons." *The Quarterly Journal of Economics* (November):1127–61.

Henisz, Witold J., Sinziana Dorobantu, and Lite J. Nartey. 2014. Spinning Gold: The Financial Returns to Stakeholder Engagement. *Strategic Management Journal*, 35(12): 1727-1748

Hennessey, J Thomas. 1998. "Reinventing" Government: Does Leadership Make the Difference? *Public Administration Review*, 58(6): 522–532.

Inoue, Carlos. 2019. Election Cycles and Organizations: How Politics Shapes the Performance of State-owned Enterprises over Time. *Administrative Science Quarterly*, 1–33.

Kivleniece, Ilze, and Bertrand V. Quelin. 2012. Creating and Capturing Value in Public-Private Ties: A Private Actor's Perspective. *Academy of Management Review*, 37(2): 272-299.

Klein, Peter G., Joseph T. Mahoney, Anita M. McGahan, and Christos N. Pitelis. Forthcoming. Organizational Governance Adaptation: Who is In, Who is Out, and Who Gets What. *Academy of Management Review*.

Krasner, Stephen D., and Thomas Risse. 2014. External Actors, State-Building, and Service Provision in Areas of Limited Statehood: Introduction. *Governance*, 27(4): 545–567.

Kroeger, Arne, and Christina Weber. 2014. Developing a Conceptual Framework for Comparing Social Value Creation. *Academy of Management Review*, 39(4): 513-540.

Levin, Jonathan and Steven Tadelis. 2010. Contracting for Government Services: Theory and Evidence from U.S. Cities. *Journal of Industrial Economics* LVIII(3):507–41.

Lindau, Luis Antonio, Dario Hidalgo, and Daniela Facchini. 2010. Curitiba, the Cradle of Bus Rapid Transit. *Built Environment*, 36(3): 350-370.

Lynn Jr., Laurence E., Carolyn Heinrich, and Carolyn J. Hill. 2000. Studying Governance and Public Management: Challenges and Prospects. *Journal of Public Administration Research and Theory*, 10(2): 233-262.

Mahoney, Joseph T., Anita M. McGahan, and Christos N. Pitelis. 2009. The Interdependence of Private and Public Interests. *Organization Science*, 20(6): 1034-1052.

Marx, Axel, and Adrian Dusa. 2011. Crisp-Set Qualitative Comparative Analysis (csQCA): Contradictions and Consistency Benchmarks for Model Specification. *Methodological Innovations Online*, 6(2):103-148.

McDonald, Robert E.. 2007. An Investigation of Innovation in Nonprofit Organizations: The Role of Organizational Mission. *Nonprofit and Voluntary Sector Quarterly*, 36(2): 256-281.

Miller, Gary J., and Andrew B. Whitford. 2016. *Above Politics*. Cambridge: Cambridge University Press.

Misangyi, Vilmos F. 2016. Institutional complexity and the meaning of loose coupling: Connecting institutional sayings and (not) doings. *Strategic Organization*, 14(4): 407-440.

Misangyi, Vilmos F., and Abhijith G. Acharya. 2014. Substitutes or Complements? A Configurational Examination of Corporate Governance Mechanisms. *Academy of Management Journal*, 57(6): 1681–1705.

Misangyi, Vilmos F., Thomas Greckhamer, Santi Furnari, Peer C. Fiss, Donal Crilly, and Ruth Aguilera. 2017. Embracing Causal Complexity: The Emergence of a Neo-Configurational Perspective. *Journal of Management*, 43(1): 255–282.

Moore, Mark H.. 1995. *Creating Public Value: Strategic Management in Government*. Cambridge: Harvard University Press.

Oakleaf, Megan. 2009. Using Rubrics to Assess Information Literacy: An Examination of Methodology and Interrater Reliability. *Journal of the American Society for Information Science and Technology*, 60(5): 969-983.

Osborne, Stephen P. 2006. The New Public Governance? *Public Management Review*, 8(3): 377-387.

Pandey, Sanjay K., and Patrick G. Scott. 2002. Red Tape: A Review and Assessment of Concepts and Measures. *Journal of Public Administration Research and Theory*, 12(4): 553-580.

Peters, B. Guy. 1998. Manage Horizontal Government: The Politics of Coordination. *Public Administration* 76(2):295–311.

Provan, Keith G., H. Brinton Milward. 2001. Do Networks Really Work? A Framework for Evaluating Public-Sector Organizational Networks. *Public Administration Review*, 61(4): 414-423.

Quelin, Bertrand V., Ilze Kivleniece, Sérgio G. Lazzarini. 2017. Public-Private Collaboration, Hybridity and Social Value: Towards New Theoretical Perspectives. *Journal of Management Studies*, 54(6): 763-792.

Raab, Jörg, Remco S. Mannak, and Bart Cambre. 2013. Combining Structure, Governance, and Context: A Configurational Approach to Network Effectiveness. *Journal of Public Administration Research and Theory*, 25, 479–511.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2006. Set Relations in Social Research: Evaluating Their Consistency and Coverage. *Political Analysis*, 14(3): 291-310.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: Chicago

University Press.

Ragin, Charles, Sarah Ilene Strand, Claude Rubinson, Kriss Drass, and Sean Davey. 2008. *USER'S GUIDE TO Fuzzy-Set / Qualitative Comparative Analysis*. (September):1–91.

Rangan, Subranian, Ramina Samii, and Luk N. Van Wassenhove. 2006. Constructive Partnerships: When Alliances Between Private Firms and Public Actors Can Enable Creative Strategies. *Academy of Management Review*, 31(3): 738-751.

Rihoux, Benoît, and Charles C. Ragin. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks: SAGE.

Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Science. A Guide to Qualitative Comparative Analysis*. New York: Cambridge University Press.

Sedgwick, Donna. 2017. Building Collaboration: Examining the Relationship Between Collaborative Processes and Activities. *Journal of Public Administration Research and Theory*, 27(2): 236–252.

Stoker, Gerry. 2006. Public Value Management: A New Narrative for Networked Governance? *The American Review of Public Administration*, 36(1): 41-57.

Suarez, David F.. 2011. Collaboration and Professionalisation: The Contours of Public Sector Funding for Nonprofit Organizations. *Journal of Public Administration Research and Theory*, 21(2): 307–326.

Tendler, Judith. 1997. *Good government in the tropics*. Baltimore: Johns Hopkins University Press.

Thomann, Eva, Nadine van Engen, and Lars Tummers. 2018. The Necessity of Discretion: A Behavioral Evaluation of Bottom-Up Implementation Theory. *Journal of Public Administration Research and Theory*, 28(4): 583–601.

Thiem, Alrik., & Baumgartner, Michael. 2016. Back to Square One: A Reply to Munck, Paine, and Schneider. *Comparative Political Studies*, *49*(6), 801–806.

Thiem, Alrik, & Dusa, Adrian. 2013. Boolean Minimization in Social Science Research: A Review of Current Software for Qualitative Comparative Analysis (QCA). *Social Science Computer Review*, *31*(4), 505–521.

Valero, Jesus N., and Hee Soun, Jang. 2016. The Role of Nonprofit Organizations in Homeless Policy Networks: A Research Note. *Cityscape*, 18(2): 151.

Villani, Elisa, Luciano Greco, and Nelson Phillips. "Understanding Value Creation in Public-Private Partnerships: A Comparative Case Study." Journal of Management Studies 54.6 (2017): 876-905.

Wagemann, Claudius and Carsten Q. Schneider. 2010. Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets. *Comparative Sociology* 9(3):397–418.

Wang, Weijie. 2015. Exploring the Determinants of Network Effectiveness: The Case of Neighborhood Governance Networks in Beijing. *Journal of Public Administration Research and Theory,* 26(2): 375–388.

Witesman, Eva M. and Sergio Fernandez. 2013. "Government Contracts With Private Organizations: Are There Differences Between Nonprofits and For-Profits?" Nonprofit and Voluntary Sector Quarterly 42(4):689–715.

Table 1. Description of the Cases and their Performance Outcomes

Sector/location	Cases with high performance	Cases with low performance
Education, Brazil	(EBh) "Sobral Education Program": Initiative implemented in 2001 to improve the quality of primary education in the municipality of Sobral.	(EBl) Project to implement new IT devices in schools aimed at improving learning and digital inclusion.
Education, India	(EIh) "Andhra Pradesh Teacher Performance Pay Initiative": this project, implemented between 2004 and 2007, sought to improve the quality of primary education through incentives (pay for performance) to teachers	(EIl) Program to improve education indicators (grades, dropouts), and healthcare indicators (nutritional needs).
Education, South Africa	(ESh) "KwaZulu-Natal - Siyakha Nentsha": school-based life-skills program for teenagers, to mitigate the threats of HIV/AIDS in the KwaZulu-Natal province, between 2008 and 2012.	(ESl) The program provided refurbished computers to schools, funding all costs of connection. The initiative involved the establishment of an operational cyberlab at the school to be used for education and the development of ICT skills.
Bureaucratic Services, Brazil	(BBh) "São Paulo Poupatempo": one-stop shop launched in 1997 to consolidate several services and bureaucratic processes involved in the issuance of official documents to citizens.	(BBl) The program started in 2011 aiming at increasing efficiency and transparency in the inspection of buildings and facilities by public agents.
Bureaucratic Services, India	(BIh) "Hyderabad e-Seva": Roll-out of electronic kiosks, launched in 1997, to support service transactions in both rural and urban areas. Kiosks offer public services and facilitate transactions with private firms (payment of phone bills, for example).	(BIl) In this project, internet and computer services were offered to members of a village community using kiosks, which were established as a self-sustained business with fees levied for various services.
Bureaucratic Services/ Urban Planning, South Africa	(BSh) "Polokwane Settlement Program": Starting in 2004, this initiative involved the coordination and centralization of multiple services related to housing, water and sanitation, electricity, health, and education. The project involved the relocation of citizens from slum dwellings to a new settlement, with formal home ownership.	(BSl) The program aimed to centralize municipal billing databases, replacing multiple, disparate IT systems, to improve the accuracy and completeness of the billing and invoicing processes, as well as improving collection and service quality.
Public Transport, Brazil	(TBh) "Curitiba BRT System": A bus rapid transport system (BRT) with capacity to accommodate more passengers at higher speeds. The busways were structured in a corridor format, crossing through central areas of the city. The project started in the 1970s but was substantially expanded thereafter.	(TBl) The program built dedicated lanes for bicycles across a large city, complemented by comprehensive education policies to control traffic and reduce speed limits.
Public Transport, India	(TIh) "New Delhi Metro": metro system providing a cost-effective and environmentally friendly alternative to existing public transport. The metro, whose construction started in 1995, eased the growing pressure on the existing, over-crowded road network.	(TIl) Launched in 2006, the program aimed to build a bus rapid transport system to upgrade the dilapidated bus infrastructure, to alleviate over-crowded roads and to complement a new metro system.
Public Transport, South Africa	(TSh) "Cape Town BRT System": creation of a bus rapid transport system, staring in 2007, in part to support the FIFA World Cup Finals, but also to provide safe, reliable and affordable transport to citizens.	(TSl) Project with the objective of providing an integrated transport system to support a large city that was about to host a large international event.
Urban Planning/ Bureaucratic Services, Brazil	(UBh) "Osasco Land Titling Regularization": Starting in 2005, the program assigned housing property rights to disadvantaged populations through the coordinated distribution of formal land titles in illegally-occupied urban areas.	(UBl) Urban planning project to transfer all city electricity cables underground and remove the electricity poles. Project goals included reducing maintenance costs and improving the urban landscape of the city.
Urban Planning, India	(UIh) "Hyderabad Metro Water Supply and Sewerage Board": The program improved the overall quality of water supply and sewerage treatment in a fast-growing Indian city. Initiated in 1989, the program was implemented through the 1990s.	(UIl) Project to provide affordable and appropriate housing to slum residents through a rehousing program intended to improves living conditions.
Urban Planning, South Africa	(USh) "eThekwini Water and Sanitation": The initiative brought potable water and sanitation facilities to a low-income metropolitan area that had practically none. The program started in the mid-1990s.	(USl) Initiative to replace the old infrastructure for electricity services. The goal was to reduce energy usage through more accurate metering and billing operations.

* For the cases with low performance, we do not identify the name of the project or specific location where it was implemented. This was a condition set by some interviewees to release confidential information.

Table 2. Configurations Consistent with High Performance

Conditions	"Internal Engagement Path"	"External Engagement Path"	
	1a	1b	1c
Public Operational Capacity	■	■	■
Collaboration with For-Profit Organizations		●	●
Collaboration with Nonprofit Organizations	⊗	●	●
Collaboration between Public Units	●		●
Stakeholder Orientation	●	●	
Consistency	0.91	0.91	0.90
Raw Coverage	0.49	0.49	0.44
Unique Coverage	0.21	0.07	0.02
Number of Cases	6	6	5
Codes of the High Performance Cases (see Table 1)	BBh, UBh, BIh, TIh, UIh, BSh	EBh, TBh, ESh, USh, TSh	TBh, USh, TSh, EIh
Exemplary Cases	New Delhi Metro, India (TIh)	KwaZulu Natal Siyakha Nentsha, South Africa (ESh)	Curitiba BRT System, Brazil (TBh)
Overall solution consistency	**0.91**		
Overall solution coverage	**0.72**		

Notes: Central conditions are represented by "●" (presence) and "⊗" (absence); contributing conditions by "●" (presence) and "⊗" (absence); necessary central conditions are represented by "■" (presence) and necessary contributing conditions are represented by "■" (presence). Blank spaces indicate a "don't care" condition; that is, the condition is not relevant to that particular configuration. Minimum thresholds used in the analysis are consistency of 0.8 and frequency of one case per configuration. Notice that a case might be present in more than one solution simultaneously. The number of cases in configuration 1b is higher than the number of reported high performance cases because this solution includes what the QCA literature refers to as a contradiction.

Table 3. Configurations Consistent with Low Performance

Conditions	2a	2b	2c	2d
Public Operational Capacity	⊗	⊗	●	⊗
Collaboration with For-Profit Organizations		⊗	⊗	●
Collaboration with Nonprofit Organizations	⊗	●	⊗	●
Collaboration between Public Units	⊗		⊗	●
Stakeholder Orientation	⊗	⊗	●	●
Consistency	1.00	1.00	0.82	1.00
Raw Coverage	0.48	0.24	0.31	0.21
Unique Coverage	0.28	0.07	0.14	0.04
Number of Cases	5	2	2	1
Codes of the Low Performance Cases (see Table 1)	BBl, UBl, UIl, USl, TSl	TIl, ESl	EBl, EIl	TBl
Overall solution consistency	**0.92**			
Overall solution coverage	**0.76**			

Notes: See Table 2.

APPENDIX

SUPPLEMENTARY TABLES AND ANALYSES

Table A1. Description of the Cases and their Performance Outcomes

Sector and location	Cases with high performance		Cases with low performance	
	Description	Why it is a case of high performance	Description*	Why it is a case of low performance
Education, Brazil	"Sobral Education Program": Initiative implemented in 2001 to improve the quality of primary education in the municipality of Sobral.	A study combined propensity score matching and differences-in-differences techniques to estimate the improvement of student's grades (Rocha, Komatsu, and Menezes Filho, 2015). The proportion of students falling behind their regular school year decreased from 57.5% to 2% between 2000 and 2014. Between 2005 and 2011, the outcomes of standardized learning assessment tests showed that students improved their grades in Portuguese and Mathematics by 24,7% and 36.4%, respectively.	Project to implement new IT devices in schools aimed at improving learning and digital inclusion.	Qualitative evidence showed that teachers were unable to use the IT devices without permanent support of IT personnel inside their classrooms. Teachers reported that they spent more time trying to operate devices rather than teaching. The IT devices were discontinued a few years after their implementation, and the equipment was considered obsolete three years after acquisition.
Education, India	"Andhra Pradesh Teacher Performance Pay Initiative": the purpose of the project, implemented between 2004 and 2007, was to improve the quality of primary education through incentives (pay for performance) to teachers	Social impact was assessed by standardized learning tests. To build the treated and control groups, the assessment study used a randomized controlled trial design (Muralidharan & Sundararaman, 2011). At the end of 2 years of the program, students in incentive schools performed significantly better than those in control schools by 22% (0.28 standard deviations) and 13% (0.16 s.d.) in math and language tests, respectively(Muralidharan & Sundararaman, 2011)(Muralidharan & Sundararaman, 2011)(Muralidharan & Sundararaman, 2011)(Muralidharan & Sundararaman, 2011). Moreover, the incentive program was found to be cost effective compared to regular bonuses paid to teachers in the short run.	Program to improve education indicators (grades, dropouts), and healthcare indicators (nutritional needs).	In 82.23% of the centers, there was scarcity of equipment like weighing machines, education kits, toys, and others; in 44% centers, none of the children enrolled were found attending pre-schools; 96% of the workers at the centers complained about not receiving payments on time; only 2.89% of the workers had special training.
Education, South Africa	"KwaZulu-Natal - Siyakha Nentsha": school-based life-skills program for teenagers, to mitigate the threats of HIV/AIDS in the KwaZulu-Natal province, between 2008 and 2012.	Through an RCT study at the classroom level, three different sets of programs were implemented: (1) a control group received standard life orientation; (2) a treatment group received an enhanced package of education focused on social skills; and (3) another treatment group received the same enhanced package plus education to develop financial skills. Compared to the control group (1), 40% to 60% of the boys reported a safer sexual behavior and 90% of the boys reported an improvement on their knowledge on HIV/AIDS. At the same time 80% of the girls improved their knowledge about the concept of savings and 50% of the girls improved their cognitive abilities. The results suggest a more responsible sexual behavior as well as higher propensity to save money and other positive behaviors among the treated teenagers (Hallman et al, 2012).	The program provided refurbished computers to schools, funding all costs of connection. The initiative involved the establishment of an operational cyberlab at the school to be used for education and the development of ICT skills.	The Cyberlab is not operational anymore (it became inactive soon after the very public launch). The Cyberlab was closed down because no dedicated personnel were made available to run the initiative as promised. Furthermore, no internet connection and training were implemented. The computers were not used and the technology rapidly became obsolete.

Sector and location	Cases with high performance		Cases with low performance	
	Description	Why it is a case of high performance	Description*	Why it is a case of low performance
Bureaucratic Services, Brazil	"São Paulo Poupatempo": one-stop shop launched in 1997 to consolidate several services and bureaucratic processes involved in the issuance of official documents to citizens.	A study compared the time spent by individuals issuing documents in cities with and without Poupatempo (Fredriksson, 2015). Using differences-in-differences analysis (i.e. changes in outcomes in the two groups before and after the introduction of Poupatempo), the study found a 29% reduction in the time spent obtaining a driver's license compared to the regular service provided by a specialized department of the public bureaucracy in the period 2008-2010.	The program started in 2011 aiming at increasing efficiency and transparency in the inspection of buildings and facilities by public agents.	The initial investment was around US$ 11.66 million, and the program consisted in the use of tablets and mini-printers for support inspection routines activities in the field. The program aimed at increasing efficiency and transparency in the bureaucratic process through the adoption of IT solutions. The government was required to make large upfront investments in IT infrastructure, devices, and customized software. Despite these efforts, the program faced several flaws not only in terms of lack of integration between software procedures and bureaucratic routines in the field, but also in terms of inappropriate IT infrastructure in the city to support operations. Months after its launch, the program was shut down due to failures in execution and implementation.
Bureaucratic Services, India	"Hyderabad e-Seva": Roll-out of electronic kiosks, launched in 1997, to support service transactions in both rural and urban areas. Kiosks offer public services and facilitate transactions with private firms (payment of phone bills, for example).	A survey study showed that the program benefitted citizens through reduced travel costs and waiting times vis-à-vis the traditional manual service, which was used as a comparison (Bhatnagar et al, 2007). For instance, waiting times at manual service centers were on average 32.9 minutes, compared to 14.5 at the e-kiosks, thus reducing waiting time by 44% relatively to the comparison group.	In this project, internet and computer services were offered to members of a village community using kiosks, which were established as a self-sustained business with fees levied for various services.	Kiosks were managed by two sets of operators. In some kiosks, a private entrepreneur was in charge, while in others they were run by an NGO. However, after more than three years of operation, most of the self-employed entrepreneurs had closed down their kiosks (by 2005, 29 of the 36 kiosks were shut down). Kiosks suffered from a host of factors including the lack of adequate infrastructure and poor package of services offered to the communities.
Bureaucratic Services/ Urban Planning, South Africa	"Polokwane Settlement Program": Starting in 2004, this initiative involved the coordination and centralization of multiple services related to housing, water and sanitation, electricity, health, and education. The project involved the relocation of citizens from slum dwellings to a new settlement, with formal home ownership.	The new settlements brought tangible improvements to the dwelling structures, improved access to services (electricity, water, health), and lower crime rates. The population attained benefits from formal land ownership, including the raised likelihood of receiving loans from banks and higher investment in home upgrades. These gains were assessed in comparison to the previous settlements that lacked consolidated services (thus serving as a comparison group). The majority of the outcome indicators increased by more than 10% (Martinez, Legovini, Krishnan, and Coville, 2011).	The program aimed to centralize municipal billing databases, replacing multiple, disparate IT systems, to improve the accuracy and completeness of the billing and invoicing processes, as well as improving collection and service quality.	The implementation took twice the initial expected time. Problems in the consolidation of the new system led to inflated water and electricity readings, and 41,000 disconnections. As consequence, there was a large number of customer complaints. These problems persisted over time after the initial implementation of the project.
Public Transport, Brazil	"Curitiba BRT System": A Bus Rapid Transport system (BRT) with capacity to accommodate more passengers at higher speeds. The busways were structured in a corridor format, crossing through central areas of the city. The project started in the 1970s but was substantially expanded thereafter.	The busways were structured in a corridor format, crossing through central areas of the city. The project started in the 1970s but was substantially expanded thereafter. Compared to regular bus systems, the new BRT system achieved higher average speeds, greater passenger capacity, and improved frequency. A study in 2009, for instance, found that the average speed was 49% higher in the BRT system compared to regular bus lines, which were used as a comparison transport mode (NTU, 2009).	"Cycle lanes": The program built dedicated lanes for bicycles across a large city, complemented by comprehensive education policies to control traffic and reduce speed limits.	The cycle lanes suffered from design problems, construction problems and poor planning and coordination with existing transport infrastructure. Usage and satisfaction levels remained very low.
Public Transport, India	"New Delhi Metro": metro system providing a cost-effective and environmentally friendly alternative to existing public transport. The metro, whose construction started in 1995, eased the growing pressure on	The metro system reduced the number of vehicles on the streets. Furthermore, a simple comparison of the project to the Kolkata Metro, another large Indian city, showed that Delhi Metro attained a cost per kilometer 62.9% lower and a return on investment 12.4% higher (Ramachandran, 2012).	Launched in 2006, the program aimed to build a Bus Rapid Transport system to upgrade the dilapidated bus infrastructure, to alleviate over-crowded roads and to	Average travel times actually increased (by 15.7 minutes), and the project suffered a financial loss of US$ 36 million per year. A survey indicated that citizen satisfaction decreased from 3.53 (considered in the survey as "average-good") to 2.54 (considered as "bad-average").

Sector and location	Cases with high performance		Cases with low performance	
	Description	**Why it is a case of high performance**	**Description***	**Why it is a case of low performance**
	the existing, over-crowded road network.		complement a new metro system.	
Public Transport, South Africa	"Cape Town BRT System": creation of a bus rapid transport system, in part to support the FIFA World Cup Finals, but also to provide safe, reliable and affordable transport to citizens. Construction started in 2007.	Although no existing impact assessment was found, we used a survey measuring citizen feedback on the BRT conducted in 2014. The survey included questions about previously available transport alternatives as well as comparisons with current transport options. The BRT scored higher than the alternatives on the many features measured, including value for money, quality, punctuality, comfort, safety, and assessments of whether citizens were proud of the new system. The highest scores were found for the indicators coding quality and perceptions of whether citizens were proud with the new system; these scores were 87.5% and 84.8% above the scores of the other transport options.	Project with the objective of providing an integrated transport system to support a large city that was about to host a large international event.	Between 2007 and 2014, more than US$ 181.5 million were spent on trying to set up the system. The BRT lanes are now misused with minibus taxis. Since the beginning, the 25 state-of-art buses bought for the price of US$ 8.6 million remain largely idle. A forensic task force recommended a full-scale investigation of the project.
Urban Planning/ Bureaucratic Services, Brazil	"Osasco Land Titling Regularization": Starting in 2005, the program assigned housing property rights to disadvantaged populations through the coordinated distribution of formal land titles in illegally-occupied urban areas.	The local government upgraded the Housing and Urban Development Department through new IT systems, georeferenced databases, and new skilled staff. By 2012, 16,944 settlements received formal land titles and in 2007 the city won the award "Selo Cidade Cidadã" ("Citizen-oriented City Certification"), offered by the Brazilian Congress. An impact study applied matching techniques and differences-in-differences estimation to examine changes in the labor supply (hours worked) of individuals in the target communities, compared to control groups. The evaluation found that lad title program increased labor supply by 80%, especially in the poorest segments. The results indicated that individuals could allocate more time to work instead of spending time protecting their belongings (Moura, Piza, Poplawski-Ribeiro, 2011).	Urban planning project to transfer all city electricity cables underground and remove the electricity poles. Project goals included reducing maintenance costs and improving the urban landscape of the city.	While the initial project set an initial target of 250 kilometers of cable to be buried in each year and an estimated completion time of no longer than five years, project execution was well below target and, by the end of 2016, conclusion was not expected within 24 years.
Urban Planning, India	"Hyderabad Metro Water Supply and Sewerage Board": The program improved the overall quality of water supply and sewerage treatment in a fast-growing Indian city. Initiated in 1989, the program was implemented through the 1990s.	An evaluation of the program before and after it was implemented pointed to the reduction in time spent on complaint redressals. Other benefits included improved accessibility, affordability and simplification of the process of attaining new plumbing connections. Compared to Chennai, the Hyderabad service was found to be 240% cheaper and less likely to encounter pipe breakages (Kamalanathan, nd).	Project to provide affordable and appropriate housing to slum residents through a rehousing program that improves living conditions.	The implementation of the project began seven years after it was approved. The project was sanctioned in 2004 but no important steps were taken for its effective implementation, resulting in a delay of over seven years. The cost of the project almost tripled from the initial proposal, and several phases and procedures were never accomplished as planned. A subsequent assessment found several irregularities and failures, including the appointment of a consultant without a transparent bidding process, incomplete identification of eligible slum dwellers, and failure to acquire private land necessary to implement the project.
Urban Planning, South Africa	"eThekwini Water and Sanitation": The initiative brought potable water and sanitation facilities to a low-income metropolitan area that had practically none. The program started in the mid 1990s.	An assessment of the outcomes of the project by the simple comparison of before and after it was implemented indicated an increase in total water-supply connections and it was the only municipality out of 14 in the same area that was considered without water vulnerability. In fact, in 2015, eThekwini presented the lowest level of the vulnerability index, which is 42% lower than the average index in the area (Mussa, 2015).	Initiative to replace the old infrastructure for electricity services. The goal was to reduce energy usage through more accurate metering and billing operations.	Eventually, the municipal government terminated the contract to implement this initiative due to economic and social losses. Citizens complained about low service quality, installation problems, and electrical damage brought by the new meters.

* For the cases of failure, we do not identify the name of the project or specific location where it was implemented. This was a condition set by some interviewees to release confidential information. Also, two cases in Brazil and South Africa mixed elements of bureaucratic services and urban development; they were classified as spanning these two areas

Table A2. Rubrics Used to Measure Outcomes and Attributes

	1	2	3	4	5
Public operational capacity "To what extent does the leader have a clear vision for the initiative and the ability to clearly articulate its goals?"	The leader shows no clear understanding of the vision behind the initiative and is unable to effectively communicate its goals.	The leader has some understanding of the vision behind the initiative, but is unclear about its goals.	The leader understands the vision behind the initiative but is unable to clearly articulate its goals.	The leader has a clear vision for the initiative and can talk about its goals.	The leader has a strong vision for the initiative and is able to clearly and enthusiastically articulate its goals.
"To what extent are employees selected based on rigorous tests with objective criteria, rather than political appointment?"	Most employees are hired through an informal selection process. In many cases, they are appointed by politicians.	Most employees are hired through an informal selection process. However, for some select positions, employees are appointed based on technical criteria.	There is a formal selection process. However, the criteria are not always accurately and explicitly defined, so it is unclear how rigorous and objective the final result is.	There is a formal, competitive selection process with some broad technical criteria. It is unlikely that external parties significantly influence recruitment decisions.	Most employees are hired through a formal, competitive selection process. Moreover, the selection process uses technical criteria that are accurately and explicitly defined. No external influence is perceived.
"To what extent is the initiative's implementation effectively project-managed, e.g. through regular monitoring and reporting of key performance indicators (KPIs), budget controls and time management?"	Implementation is not managed. There is no dedicated project management office and effective monitoring does not take place. There is no regular oversight and reporting of budgets or timelines.	There is some monitoring of budgets and timelines, but key performance indicators are not well defined. Reporting is ad-hoc and quality and efficiency are not measured. Issues are dealt with when they arise and this can mean that implementation is delayed or stopped while the solutions are sought.	Individuals have project management responsibilities, but this is in parallel to their other responsibilities. Key performance indicators are defined and collected, but processes to act upon those reports are weak. Budgets and timelines are also tracked, but not effectively managed when things are moving off course.	A project management office exists to monitor and report on key performance indicators and to track budgets and timelines. The tracking takes place, but the team is not empowered with processes to resolve all issues.	A dedicated project management office regularly monitors and reports on key performance indicators and ensures that budgets and timelines are met. Any issues are raised before they can impact performance and are escalated to management for resolution. The team is empowered to act upon emerging issues as they are identified.
"To what extent is the workforce supported to adapt to new processes and systems?"	The workforce is left to adapt to change with no support.	Employees are sent memos to inform them about new processes and systems, but there is no further engagement to provide explanations and support.	Employees are made aware of new processes and systems in advance of their implementation. Material is provided to the employees which gives further explanations and instructions. For more complex systems, a training session is organized to instruct employees. Employees are also encouraged to ask questions of their supervisors regarding aspects they are unsure about.	A clear change-management process is implemented which provides employees with information about upcoming changes to processes and systems in advance, provides training sessions and materials to prepare them for new systems and encourages them to ask questions of their supervisors regarding aspects they are unsure about.	A dedicated change-management team is responsible for implementing a clear change-management process which includes regular training sessions and workshops to ensure that the systems can be adapted to meet user needs more effectively. Continuous feedback processes are instated to ensure that questions are answered promptly and ensuring that the systems themselves are responsive to feedback from the users.
"To what extent is implementation supported by monitoring and enforcement apparatus to prevent corruption (e.g. anti-corruption rules with high likelihood of punishment)?"	There are no mechanisms to identify potential corruption or to effectively enforce against it if it is uncovered.	There is a mechanism to identify potential corruption. However, the monitoring system is not effective and where corruption is uncovered, enforcement is weak.	There are well-established mechanisms to detect corruption. However, enforcement is weak and punishment for deviants is unlikely.	There are well-established mechanisms to detect corruption. In some cases, deviants are punished. Minor-level corruption is sometimes overlooked, and incidents involving senior staff are rarely enforced.	There are well-established mechanisms to detect and avoid corruption. Full enforcement and punishment are expected, irrespective of the scale of the offense and seniority of the deviant.

Collaboration with for-profit organizations "To what extent does the unit responsible for the project collaborate effectively with the private (for-profit) sector?"	The initiative functions completely independently of the private sector.	The initiative functions largely independently of organizations from the private sector, but does include them in some support activities.	The initiative functions largely independently of organizations in the private sector with regard to core operations. However, support activities tend to have some involvement from them.	The initiative controls its core operations but is dependent on private sector organizations for many shared activities.	The initiative's operations are deeply connected with the private sector. The initiative is in effect a partnership characterized by mutual interdependence.
"To what extent do the employees of the initiative come from private (for-profit) sector organizations?"	No employees of the initiative come from the private sector.	Migration from the private sector is rare and it is not generally seen as an option for employees in that sector.	Migrations are not common. However, employees from the private sector do see it as a feasible opportunity and the initiative's management is open to the idea for certain positions.	The initiative's management encourages migration from the private sector. However, this rarely occurs in top management positions.	The initiative's management actively encourages migration from the private sector. They have created incentives to attract such recruits and have processes to target individuals from this sector.
Collaboration with nonprofit organizations "To what extent does the unit responsible for the project collaborate effectively with (nonprofit) organizations of civil society?"	The initiative functions completely independently of civil society.	The initiative functions largely independently of organizations from civil society, but does include them in some support activities.	The initiative functions largely independently of organizations in civil society with regard to core operations. However, support activities tend to have some involvement from them.	The initiative controls its core operations but is dependent on civil-society organizations for many shared activities.	The initiative's operations are deeply connected with civil society. The initiative is in effect a partnership characterized by mutual interdependence.
"To what extent do the employees of the initiative come from (nonprofit) civil society organizations?"	No employees of the initiative come from civil society.	Migration from civil society organizations is rare and it is not generally seen as an option for employees in that sector.	Migrations are not common. However, employees from civil society organizations do see it as a feasible opportunity and the initiative's management is open to the idea for certain positions.	The initiative's management encourages migration from civil society organizations. However, this rarely occurs in top management positions.	The initiative's management actively encourages migration from civil society organizations. They have created incentives to attract such recruits and have processes to target individuals from this sector.
Collaboration between public units "To what extent does the unit responsible for the project collaborate with other organizations in the public sector?"	The initiative functions completely independently of other organizations in the public sector.	The initiative functions largely independently of other organizations in the public sector but does share some support activities with other public organizations.	The initiative functions largely independently of other organizations in the public sector with regards to core operations. However, support activities tend to be shared with other public units.	The initiative controls its core operations but is dependent on other public organizations for many shared activities.	The initiative's operations are connected with those of other public organizations. The initiative is in effect co-managed with other related initiatives, in a system characterized by mutual interdependence.
"To what extent do the employees of the initiative come from other parts of the public sector?"	No employees of the initiative come from other parts of the public sector.	Migration from other public sector organizations is rare and it is not generally seen as an option for most public service employees.	Migrations are not common. However, employees do see it as a feasible opportunity and the management is open to the idea for certain positions.	The initiative's management encourages migration from other public service organizations. However, this rarely occurs in top management positions.	The initiative's management actively encourages migration from other public service organizations. They have created incentives to attract such recruits and have processes to target individuals from other public service organizations.
Stakeholder orientation "To what extent does project leadership allow ideas and solutions to come from different parts of the organization and stakeholders?"	The leader requires all strategic and operational decisions to be taken by top management, with his personal approval, and without the involvement of other parts of the workforce or other stakeholders.	The leader requires all strategic decisions to be taken by top management, with his personal approval, and operational decisions are also taken by top management, with limited involvement from other parts of the workforce or other stakeholders.	Top leadership has ultimate responsibility for strategic and operational decision making, but space is given for middle and senior management to provide their input to the decision-making process. The broader workforce and other stakeholders are excluded from these processes.	Processes exist to allow input from management as well as the broader workforce to be considered in decision making, for both strategic and operational issues. There is also a forum to allow stakeholders beyond the workforce to submit their thoughts for consideration.	The top leadership team actively encourages and incentivizes the input and sharing of ideas from across the workforce. Processes exist to evaluate and incorporate these ideas into the formal decision-making processes of the initiative. Mechanisms also exist to allow the exchange of ideas with stakeholders outside the workforce, and processes are in place to bring

5

	1	2	3	4	5
					these ideas into the decision-making activities of the initiative.
"To what extent is project leadership visibly connected to the performance of the initiative through proactive internal and external communication?"	The leader is not involved in any external communication and is not visibly connected to the initiative as far as the public is concerned. Internally, the leader is absent from the point of view of the workforce.	The leader rarely engages in formal communication either internally or externally.	The leader makes some announcements externally, when required, and is seen to be connected to the project. Communication with the workforce is limited.	The leader is a proactive communicator of the initiative externally, with some presence in the media, and communicates high-level developments to the workforce.	The leader has a regular media presence and a reputation that is closely tied with the performance of the initiative and takes an active interest in regularly communicating through formal and informal channels.
"To what extent are decision-making and change in the public-sector environment constrained by formal rules?" (reverse scored)	Formal rules hold absolute authority across the public sector environment and are strictly enforced by hierarchical structures. There is no room to question or change existing rules. Attempts to do so are not tolerated.	Formal rules are strictly enforced across the public sector environment. Some parts of the public sector have processes and mechanisms to adapt these rules, but these are complex and rarely used.	Formal rules form the foundation of the institutional environment, but senior administrators and politicians intermittently revisit the relevance and nature of these rules to ensure their relevance to current conditions. Public service employees are usually consulted about such changes.	Politicians and senior administrators in the public sector allow suggestions for the modification of existing rules at regular intervals. Processes and mechanisms exist to ensure that public service employees are able to influence and contribute towards the modification of these rules.	Formal rules exist but there is a continuous process to allow for modifications to these rules, in response to changing circumstances. Public service employees are able to fully engage in this process, their ideas are welcomed, encouraged and actively evaluated by politicians and senior administrators. Other stakeholders, including the broader public are also encouraged to participate in this process.
"To what extent is change in the public sector environment facilitated by norms in the workplace?"	Public service employees feel threatened by changes in their institutional environment and respond confrontationally to any attempt to modify existing processes and ways of working. Employees defend their own turf vigorously and sideline any individuals that seem like they might engender change. Public sector managers punish employees that do not fit the norm.	Public service employees generally resist changes to existing processes and ways of working. They can be convinced to perform minor changes that do not meaningfully compromise the nature of their informal norms. Employees tend to defend their own turf and public sector managers protect the status quo.	Public service employees have the capacity for change, but require incentives and support to do so. When change happens, it tends to be passively accepted rather than enthusiastically embraced. Managers in the public sector support change but are not proactive in engendering it.	Public service employees are generally open to change and enthusiastically adopt new ideas and ways of working. Public service managers are supportive of change and have processes to help employees adapt to new systems and ways of working.	Public service employees are highly motivated to start and promote changes and improvements to the processes and ways of working in the public sector. It is common for employees and management to regularly collaborate to generate ideas for continuous improvement as well as considering more transformational opportunities for improvement.
Performance outcome "To what extent is there robust evidence of positive social impact?"	The project was not implemented and/or was discontinued due to strong opposition or criticism regarding its potential outcomes and costs. In cases where the project was implemented, there is qualitative or quantitative evidence that the target populations received negative or scant impact as a result of the project.	The project was implemented but its initial targets were not met. There is some qualitative or quantitative evidence that the project was not generating positive social impact, considering its initial required investment.	There is qualitative or quantitative evidence that some positive impact occurred on some important dimensions, but results are not totally aligned (for instance, there is impact on some dimensions but not on others).	There is quantitative evidence that positive impact occurred on some important dimensions of performance, based on independent studies using comparison groups (populations not affected by the intervention).	There is strong quantitative evidence that positive impact occurred on some important dimensions of performance valued by beneficiaries, based on independent studies using robust counterfactual analysis (randomized controlled trials, matching analysis, or differences-in-differences estimation).

Table A3. Descriptions of Interviewees for Each Case

Sector and location	Cases with evidence of success (high performance)			Cases with evidence of failure (low performance)		
	Interviewee 1	Interviewee 2	Interviewee 3	Interviewee 1	Interviewee 2	Interviewee 3
Education, Brazil	Secretary of Education	Director of nonprofit	School principal	Secretary of Education	Municipal manager	School vice-principal
Education, India	Leader of local nonprofit	Vice-chairman at a private organization	Senior economist at nonprofit	Professor at a local university	Project supervision at government unit	Researcher at a public organization
Education, South Africa	Researcher specialized in education	Program manager at nonprofit	Senior researcher at nonprofit	Headmaster at a local public school	Project leader at government unit	Professor at a local university
Bureaucratic Services, Brazil	General director of public unit	Chief advisor at public unit	Researcher assessing the project	Union leader	Supervisor of municipal public services	Municipal manager
Bureaucratic Services, India	Professor at a local university	Director at a private organization	Assistant manager at a local university	Project consultant at a private organization	Director at a social entrepreneurship center	Professor at a local university
Bureaucratic Services/ Urban Planning, South Africa	Researcher at a nonprofit	Director at a national department	Public manager at a local unit	Municipal Public Committee member	Consultant	Subcontractor at a for-profit firm
Public Transport, Brazil	President of planning institute	Former Secretary of Transport and Mayor	Chief engineer in the transport division	Former Secretary of Transport	Nonprofit activist	Engineer responsible for project implementation
Public Transport, India	Expert in public transportation	Professor at a local university	Operations manager at a government unit	Principal scientist at a national laboratory	Researcher at an international university	Professor at a local university
Public Transport, South Africa	Professor at a local university	Counselor for municipal transport	Professor at a local university	Member at a local nonprofit	Member at a local nonprofit	Transport engineer in for-profit firm
Urban Planning/ Bureaucratic Services, Brazil	Director in the unit of housing and urbanism	Public employee in social services	Director of social services	Former CEO of public supply firm	Media professional and researcher	Director of infrastructure at a government unit
Urban Planning, India	Fellow at an international university	Professor at an international university	Social development consultant	Director at a leading nonprofit	Professor at a local university	Professor at a local university
Urban Planning, South Africa	Strategic executive at a municipality department	Professor at a local university	Asset manager at the municipality	Committee member at a local municipality	Business analyst at a nonprofit	Media professional

Table A4. Raw Data Matrix

Sector and location (project code in parenthesis)	Public Operat. Capacity	Collab. with For-Profit Organ.	Collab. with Nonprofit Organ.	Collab. between Public Units	Stakeholder Orient.	Performance Outcome
Bureaucratic Services, South Africa (BSh)	3.90	2.00	2.00	3.50	3.17	5.00
Urban Planning, Brazil (UBh)	3.60	2.67	1.00	4.17	4.17	5.00
Bureaucratic Services, Brazil (BBh)	4.63	2.00	2.34	4.50	3.79	5.00
Education, South Africa (ESh)	5.00	5.00	5.00	2.17	4.67	5.00
Education, Brazil (EBh)	4.87	2.84	4.00	2.17	3.83	5.00
Education, India (EIh)	4.33	3.67	3.67	3.83	2.42	5.00
Transportation, Brazil (TBh)	4.33	3.34	3.00	4.34	3.92	4.33
Transportation, India (TIh)	4.73	3.17	1.67	3.84	4.08	4.33
Bureaucratic Services, India (BIh)	3.93	4.67	2.00	3.67	3.25	4.33
Urban Planning, India (UIh)	4.07	2.17	1.50	3.34	2.67	4.33
Urban Planning, South Africa (USh)	5.00	3.83	3.34	3.84	4.67	3.67
Transportation, South Africa (TSh)	5.00	3.67	3.34	3.34	4.00	3.00
Transportation, South Africa (TSl)	1.20	4.67	1.17	2.17	2.00	2.00
Transportation, Brazil (TBl)	2.50	3.17	3.67	3.17	3.21	1.67
Urban Planning, India (UIl)	1.13	2.17	1.50	1.34	2.09	1.67
Bureaucratic Services, South Africa (BSl)	3.10	3.17	1.33	3.17	2.50	1.67
Education, India (EIl)	4.40	1.00	2.00	2.00	3.00	1.67
Bureaucratic Services, India (BIl)	3.80	3.50	4.00	3.50	3.17	1.67
Education, South Africa (ESl)	1.50	2.50	2.84	1.34	2.29	1.33
Urban Planning, Brazil (UBl)	1.17	1.17	1.17	1.50	1.59	1.33
Education, Brazil (EBl)	2.93	2.34	1.00	2.17	3.25	1.33
Transportation, India (TIl)	2.00	2.00	3.33	4.00	1.67	1.00
Bureaucratic Services, Brazil (BBl)	1.73	2.34	1.00	2.34	2.50	1.00
Urban Planning, South Africa (USl)	1.57	5.00	1.00	1.00	2.25	1.00

Note: The table reports average scores for each condition and the performance outcome based on coded interviews (three distinct raters per case). Each coded condition is also a composite measure involving multiple items representing the underlying construct. The labels in the first column include project codes in parenthesis following Table 1.

Table A5. Reliability Indicators of the Measures used in the Cases

Sector and location	Cases with high performance			Cases with low performance		
	Cronbach Alpha	Kappa agreement test (*p* value)*	Intraclass correlation (ICC)	Cronbach Alpha	Kappa agreement test (*p* value)*	Intraclass correlation (ICC)
Education, Brazil	0.839	< 0.000	0.834	0.905	< 0.000	0.900
Education, India	0.895	< 0.000	0.895	0.949	< 0.000	0.949
Education, South Africa	0.931	< 0.000	0.931	0.810	< 0.000	0.806
Bureaucratic Services, Brazil	0.819	< 0.000	0.839	0.871	< 0.000	0.868
Bureaucratic Services, India	0.853	< 0.000	0.853	0.822	< 0.000	0.822
Bureaucratic Services/ Urban Planning, South Africa	0.843	< 0.000	0.847	0.732	< 0.000	0.717
Public Transport, Brazil	0.781	< 0.000	0.777	0.758	< 0.013	0.771
Public Transport, India	0.776	< 0.000	0.776	0.832	< 0.000	0.831
Public Transport, South Africa	0.729	< 0.000	0.724	0.896	< 0.000	0.896
Urban Planning/ Bureaucratic Services, Brazil	0.797	< 0.004	0.818	0.831	< 0.010	0.902
Urban Planning, India	0.923	< 0.000	0.923	0.853	< 0.000	0.853
Urban Planning, South Africa	0.818	< 0.000	0.814	0.895	< 0.000	0.887

*The null hypothesis is that responses across interviewees were randomly determined.

Table A6. Calibrated Data Matrix

Sector and location	Public Operat. Capacity	Collab. with For-Profit Organ.	Collab. with Nonprofit Organ.	Collab. between Public Units	Stakeholder Orient.	Performance Outcome
Bureaucratic Services, South Africa (BSh)	1.00	0.33	0.33	0.67	0.67	1.00
Urban Planning, Brazil (UBh)	0.67	0.67	0.00	1.00	1.00	1.00
Bureaucratic Services, Brazil (BBh)	1.00	0.33	0.33	1.00	1.00	1.00
Education, South Africa (ESh)	1.00	1.00	1.00	0.33	1.00	1.00
Education, Brazil (EBh)	1.00	0.67	1.00	0.33	1.00	1.00
Education, India (EIh)	1.00	0.67	0.67	1.00	0.33	1.00
Transportation, Brazil (TBh)	1.00	0.67	0.67	1.00	1.00	1.00
Transportation, India (TIh)	1.00	0.67	0.33	1.00	1.00	1.00
Bureaucratic Services, India (BIh)	1.00	1.00	0.33	0.67	0.67	1.00
Urban Planning, India (UIh)	1.00	0.33	0.33	0.67	0.67	1.00
Urban Planning, South Africa (USh)	1.00	1.00	0.67	1.00	1.00	0.67
Transportation, South Africa (TSh)	1.00	0.67	0.67	0.67	1.00	0.67
Transportation, South Africa (TSl)	0.00	1.00	0.00	0.33	0.33	0.33
Transportation, Brazil (TBl)	0.33	0.67	0.67	0.67	0.67	0.33
Urban Planning, India (UIl)	0.00	0.33	0.33	0.33	0.33	0.33
Bureaucratic Services, South Africa (BSl)	0.67	0.67	0.33	0.67	0.33	0.33
Education, India (EIl)	1.00	0.00	0.33	0.33	0.67	0.33
Bureaucratic Services, India (BIl)	1.00	0.67	1.00	0.67	0.67	0.33
Education, South Africa (ESl)	0.33	0.33	0.67	0.33	0.33	0.33
Urban Planning, Brazil (UBl)	0.00	0.00	0.00	0.33	0.33	0.33
Education, Brazil (EBl)	0.67	0.33	0.00	0.33	0.67	0.33
Transportation, India (TIl)	0.33	0.33	0.67	1.00	0.33	0.00
Bureaucratic Services, Brazil (BBl)	0.33	0.33	0.00	0.33	0.33	0.00
Urban Planning, South Africa (USl)	0.33	1.00	0.00	0.00	0.33	0.00

Note: We applied the recoding calibration method (Emmenegger, Schraff, and Walter 2014; Schneider and Wagemann 2012; Thomann et al. 2018). Please refer to more details in Table A7 of the supplementary appendix.

Table A7. Final Calibration of Attributes

	"Fully out" (0.00)	"More out than in" (0.33)	"More in than out" (0.67)	"Fully-in" (1.00)
Public Operational Capacity	12.5%	20.8%	12.5%	54.2%
Collaboration between Public Units	4.2%	37.5%	29.2%	29.2%
Collaboration with For-Profit Organ.	8.3%	33.3%	37.5%	20.8%
Collaboration between Public Units	4.2%	37.5%	29.2%	29.2%
Stakeholder Orientation	0.0%	37.5%	29.2%	33.3%
Performance Outcome	12.5%	37.5%	8.3%	41.7%

Note: We applied the recoding calibration method (Emmenegger, Schraff, and Walter 2014; Schneider and Wagemann 2012; Thomann et al. 2018). In this method, researchers have to choose a scale to classify presence or absence, as well as intermediate situations. We opted for a four-value fuzzy scale, as follows: 0 = "fully out", 0.33 = "more out than in", 0.67 = "more in than out", and 1 = "fully in." Next, the method requires the definition of crossover points, which will then reclassify the conditions in each point of the recoded scale. We defined 2.5 in the average score of the composite measure as our crossover point; 1.25 as our "fully out" threshold point; and 3.75 as our "fully in" threshold point (i.e. for a condition to be classified as "fully in," its composite raw score from our rubrics should be at least 3.75).

We considered 2.5 as our crossover point, rather than 3, given that the composite nature of our rubrics creates the risk that a case will fall just below 3 even when some items indicate presence of the condition. Consider again the condition "collaboration with for-profit organizations" in the Sobral Education Program. Although the final raw score of this condition was slightly below 3, as explained before, our qualitative assessment of the project indicates that there was moderate engagement of some private firms providing project managers with specialized input. Therefore, the score should represent a "more in than out" (0.67) situation, thus justifying the crossover point of 2.5 instead of 3. We used a similar approach to establish the full membership and full non-membership anchors (respectively: 3.75 and 1.25 average scores). Notice therefore that our recoding method of calibration follows methodological recommendations to consider relevant information (Emmenegger et al. 2014, 7–8) and use of our in-depth knowledge of the cases (Schneider and Wagemann 2012, 34–35).

Additionally, following recent advances in QCA methodology (Thomann et al. 2018), we performed a robustness check by using an alternative technique, direct method of calibration, using the same threshold points, as well as consistency and frequency thresholds. Even though different calibrations substantially affect the configurational result (Thomann et al. 2018), the main results were quite similar to those reported in our analysis below (see Table A9 of the supplementary appendix).

Table A8. Truth Table for High Performance

Conditions					Performance Outcome (1 = High Performance)	Number of Cases	Codes of Cases (see Table 1)	Consistency
Public Operational Capacity	Collaboration with For-Profit Organ.	Collaboration with Nonprofit Organ.	Collaboration between Public Units	Stakeholder Orientation				
1	1	1	0	1	1	2	EBh, ESh	1.000
1	1	1	1	0	1	1	EIh	0.900
1	1	1	1	1	1	4	TBh, USh, TSh, BIl	0.899
1	1	0	1	1	1	3	UBh, BIh, TIh	0.895
1	0	0	1	1	1	3	BBh, UIh, BSh	0.889
0	0	1	0	0	0	1		0.795
1	1	0	1	0	0	1		0.725
1	0	0	0	1	0	2		0.723
0	0	1	1	0	0	1		0.663
0	1	1	1	1	0	1		0.663
0	1	0	0	0	0	2		0.596
0	0	0	0	0	0	3		0.595

Note: We used a consistency level of 0.800 and a frequency threshold of at least one representative case per solution. Notice that, by examining the truth table, our main results would hold even with a higher consistency threshold of 0.850. For illustration, in the second row of Table A8, there is a configuration involving the presence of public capacity, all three forms of collaboration, and the absence of stakeholder orientation. One case of high performance fits this configuration: education in India (Andhra Pradesh Teacher Performance Pay Initiative). Roughly speaking, consistency is coded as 0.900 because this particular combination has a high degree of conformity (0.900) with the high-performance outcome (Rhioux and Ragin 2009:108). The configuration in the third row includes a contradiction, that is, a case of low performance (BIl). We opted to keep the contradiction row in the logical minimization process following recommendations by the specialized QCA literature (Schneider and Wagemann 2012:122). For robustness, we also re-ran the fsQCA algorithm without the contradictory case (Thiem and Dusa 2013: 510), and our conclusions remained the same. The analysis of the truth table allowed us to identify potential gaps in the consistency scores (Bell, Filatotchev, and Aguilera 2014; Campbell, Sirmon, and Schijven 2016; Fiss 2007, 2011).

Table A9. Direct Method Calibration Results: Configurations Consistent with Evidence of High Performance

Conditions	High Performance of Public Initiatives		
	3a	**3b**	**3c**
Public Operational Capacity	■	■	■
Collaboration with For-Profit Organ.		●	•
Collaboration with Nonprofit Organ.	⊗	•	●
Collaboration between Public Units	●	⊗	•
Stakeholder Orientation	●	●	⊗
Consistency	0.86	0.97	0.86
Raw Coverage	0.45	0.19	0.14
Unique Coverage	0.35	0.10	0.04
Number of Cases	6	2	1
Codes of the Cases (See Table 1)	BBh, UBh, BIh, TIh, UIh, BSh	EBh, ESh	EIh
Overall solution consistency	**0.87**		
Overall solution coverage	**0.61**		

Notes: Central conditions are represented by "●" (presence) and "⊗" (absence); contributing conditions by "•" (presence) and "⊗" (absence); necessary central conditions are represented by "■" (presence) and necessary contributing conditions are represented by "■" (presence). Blank spaces indicate a "don't care" condition; that is, the condition is not relevant to that particular configuration. Minimum thresholds used in the analysis are consistency of 0.8 and frequency of one case per configuration.

Table A10. The Necessity of the Conditions for High Performance

	Consistency	*p*-value for consistency benchmark of		
		0.5 ("more often than not")	0.65 ("usually")	0.8 ("almost always")
Public Operational Capacity	0.91	0.003***	0.042**	0.275
Collaboration with For-Profit Organ.	0.70	0.073*	0.347	0.795
Collaboration with Nonprofit Organ.	0.58	0.387	0.787	0.981
Collaboration between Public Units	0.84	0.003***	0.042**	0.275
Stakeholder-Centered Governance	0.88	0.003***	0.042**	0.275
Collab. with Public *or* Collab. with For-Profit *or* Collab. with Nonprofit Organizations	0.95	0.000***	0.006***	0.069*

Note: *** *p* < 0.01; ** *p* < 0.05 and * *p* < 0.10.

Our necessity analysis follows recommendations in the literature (Dwivedi et al, 2018; Ragin, 2000, 2008). The main idea behind this analysis is to statistically evaluate whether each condition (or combination of conditions) is necessary to explain outcomes. For each condition, we considered three benchmark levels, namely: "more often than not" (0.5), "usually" (0.65) and "almost always" (0.8) (Ragin, 2000: 109-110). Since we have less than 30 cases, instead of following the formula using the binomial-normal approximation, we used the binomial distribution (Ragin, 2000: 112) to conduct our hypothesis tests assuming successes and failures were strictly binary considering our estimated consistency and the number of causal conditions/combinations under analysis.

Table A11. The Necessity of the Conditions for Low Performance

	Consistency	p-value for consistency benchmark of		
		0.5 ("more often than not")	0.65 ("usually")	0.8 ("almost always")
~Public Operational Capacity	0.62	0.194	0.583	0.927
~Collaboration with For-Profit Organ.	0.62	0.194	0.583	0.927
~Collaboration with Nonprofit Organ.	0.79	0.019**	0.151	0.558
~Collaboration between Public Units	0.72	0.073*	0.347	0.795
~Stakeholder-Centered Governance	0.69	0.073*	0.347	0.795
~Collab. with Public *or* ~Collab. with For-Profit *or* ~Collab. with Nonprofit Organizations	0.90	0.003***	0.042**	0.275

Note: *** *p* < 0.01; ** *p* < 0.05 and * *p* < 0.10.

Our necessity analysis follows recommendations in the literature (Dwivedi et al, 2018; Ragin, 2000, 2008). The main idea behind this analysis is to statistically evaluate whether each condition (or combination of conditions) is necessary to explain outcomes. For each condition, we considered three benchmark levels, namely: "more often than not" (0.5), "usually" (0.65) and "almost always" (0.8) (Ragin, 2000: 109-110). Since we have less than 30 cases, instead of following the formula using the binomial-normal approximation, we used the binomial distribution (Ragin, 2000: 112) to conduct our hypothesis tests assuming successes and failures were strictly binary considering our estimated consistency and the number of causal conditions/combinations under analysis.

Table A12. Frequency Analysis of Configurations Consistent with High Performance by Sector

Configuration	Sector			
	Education	Public Transport	Bureaucratic Services	Urban Planning
1a	0	1	3	2
1b	2	2	0	1
1c	1	2	0	1

Table A13. Frequency Analysis of Configurations Consistent with High Performance by Country

Configuration	Country		
	Brazil	India	South Africa
1a	2	3	1
1b	2	0	3
1c	1	1	2

Table A14. Truth Table for Low Performance

	Conditions				Performance Outcome (1 = High Performance)	Number of Cases	Codes of the Cases (see Table 1)	Consistency
Public Operational Capacity	Collaboration with For-Profit Organ.	Collaboration with Nonprofit Organ.	Collaboration between Public Units	Stakeholder Orientation				
0	0	1	0	0	0	1	ESl	1.000
0	0	1	1	0	0	1	TIl	1.000
0	1	1	1	1	0	1	TBl	1.000
0	1	0	0	0	0	2	TSl, USl	1.000
0	0	0	0	0	0	3	BBl, UBl, UIl	1.000
1	0	0	0	1	0	2	EBl, EIl	0.819
1	1	0	1	0	1	1		0.637
1	1	1	1	0	1	1		0.498
1	0	0	1	1	1	3		0.442
1	1	0	1	1	1	3		0.419
1	1	1	0	1	1	2		0.415
1	1	1	1	1	1	4		0.399

Note: We used a consistency level of 0.800 and a frequency threshold of at least one representative case per solution.

Table A15. Complementary Conditions Associated with Each Proposed Theoretical Path

Configuration	Complementary Conditions
"Internal engagement path" (1a)	• Public operational capacity helps foster internal collaborations between multiple government units. Internal collaborations increase the performance-enhancing effect of public operational capacity by helping improve internal processes as result of coordinated action within the public bureaucracy. • Stakeholder orientation increases the performance-enhancing effect of public operational capacity by improving the ability of public managers to connect with and incorporate valuable suggestions from multiple actors. The combination of internal resource mobilization and permeability to external input increases the legitimacy of the initiative. • Collaboration between multiple public units allows for the articulation of multiple resources and capabilities in the public sector when external collaborations are not necessarily present and especially when they are absent (this path, in particular, does not involve key collaborations with nonprofit organizations).
"External engagement path" (1b and 1c)	• Combination of complementary collaborations with for-profit and nonprofit organizations: the former can bring critical resources such as proprietary technical knowledge, extra funding, and efficient execution capabilities, while the latter can contribute with specialized knowledge of target beneficiaries, skills to communicate with local communities, or philanthropic resources to support interventions. The presence of mission-driven nonprofits can help balance social and profit-oriented objectives during project design and implementation, thereby increasing legitimacy. • Public operational capacity increases accountability and mitigates risk or conflict in the presence of external actors. Capable public managers foster legitimate interactions with the private sector and help reduce constraints to private engagement. • Although the presence of external collaborations allows for the incorporation of novel and valuable input, there is still the need to consider and articulate multiple contributions to the project, including contributions emanating from multiple public units. This can be done in two ways: either via superior stakeholder orientation (increasing the permeability of the focal public unit to receive additional input from other relevant areas of the public sector) or internal collaborations between public units (which may bring different ideas and resources to improve project design and adaptation).

Table A16. Other Idiosyncratic Conditions Likely Affecting the Outcomes of the High Performance Cases

Sector and location	Other case-specific conditions
Education, Brazil (EBh)	Local politicians had a particular preoccupation with projects in education and realized that improvements in this area could yield electoral benefits. Politicians tried to create a local intrinsic motivation to pursue high-quality education. For instance, a former mayor of the city declared that "those who teach do it for love, not for the salary. If you want to earn more, resign and go to a private school."
Education, India (EIh)	The project received significant support given the presence of researchers who would like to test the effect of incentives. The experiment led by a team of experts from internationally recognized academic institutions intended to confirm the relationship between incentives for teachers and schools and student performance.
Education, South Africa (ESh)	An HIV outbreak increased the effort to pursue the project. A rural area with high incidence of HIV caught the attention of scholars and nonprofit managers to improve the situation working with the local youth community.
Bureaucratic Services, Brazil (BBh)	The project started in São Paulo, a megacity and the most developed city in Brazil. Its growing complexity and disordered growth increased pressure to improve public services and modernize systems at the citizen interface.
Bureaucratic Services, India (BIh)	Traditionally, Hyderabad has been open to disruptive technological implementations. For instance, in 1999, it implemented a high-technology business center for biomedical research, the Genome Valley. It is also known as the "hotbed of e-government in India."
Bureaucratic Services/ Urban Planning, South Africa (BSh)	The end of apartheid brought a new urban policy with legislation aimed at improving the lives of slum dwellers. Rapid urbanization was taking place in cities across the country, accompanied by congestion, deteriorating environmental quality, and increasing cost of urban services. Disteneng, an informal settlement situated on the outskirts of Polokwane with a majority of black citizens, became a natural target to address those pressing issues.
Public Transport, Brazil (TBh)	The first mayor leading the project, was an architect. Thus, he was a technical politician. His background helped promote new systemic designs, reinforcing the need of innovative and integrative solutions throughout the city.
Public Transport, India (TIh)	Delhi's increasingly overcrowded bus system helped to accelerate car sales as more commuters opted to drive instead of taking public transportation. The situation required urgent action from the public services to avoid major chaos in the city.
Public Transport, South Africa	The project was triggered by a rare megaevent, the World Cup tournament. Public managers and citizens were motivated to make the event happen, and this required major improvements in the city.
Urban Planning/ Bureaucratic Services, Brazil (UBh)	Osasco is part of the metropolitan region of São Paulo, with high population density. However, different from the city of São Paulo itself, Osasco had more room to expand habitable areas. Granting land titles to slum dwellers was part of this effort.
Urban Planning, India (UIh)	The region was facing increased demand for water and a sewerage system in collapse. The situation was causing discomfort and incidence of diseases (such as cholera), thus creating growing pressure to find new solutions.
Urban Planning, South Africa (USh)	The newly formed democratic governments increasingly became more sensitive to the millennium development goals. The project, in particular, tried to address the shortcomings of a distant and hilly location, with poor public services.

References

Bell, R. Greg, Igor Filatotchev, and Ruth V. Aguilera. 2014. Corporate Governance and Investors' Perceptions of Foreign IPO Value: An Institutional Perspective. *Academy of Management Journal*, 57(1): 301–320.

Campbell, Joanna Tochman, David G. Sirmon, and Mario Schijven. 2016. Fuzzy Logic and the Market: A configurational approach to investor perceptions of acquisition announcements. *Academy of Management Journal*, 59(1): 163–187.

Dwivedi, Priyanka, Aparna Joshi, and Vilmos F. Misangyi. 2018. Gender-Inclusive Gatekeeping: How (Mostly Male) Predecessors Influence the Success of Female Ceos. *Academy of Management Journal*, 61(2): 379–404.

Emmenegger, Patrick, Dominik Schraff, and André Walter. 2014. *QCA, the Truth Table Analysuis and Large-N Survey Data: The Benefits of Calibration and the Importance of Robutness Test*. Compass Working Paper Series.

Fiss, Peer C. 2007. A Set-Theoretic Approach To Organizational Configurations. *Academy of Management Review*, 32(4): 1180–1198.

Fiss, Peer C. 2011. Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research. *Academy of Management Journal,* 54(2): 393–420.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: Chicago University Press.

Rihoux, Benoît, and Charles C. Ragin. 2009. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Thousand Oaks: SAGE.

Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Science. A Guide to Qualitative Comparative Analysis*. New York: Cambridge University Press.

Thomann, Eva, Nadine van Engen, and Lars Tummers. 2018. The Necessity of Discretion: A Behavioral Evaluation of Bottom-Up Implementation Theory. *Journal of Public Administration Research and Theory*, 28(4): 583–601.

Thiem, Alrik, & Dusa, Adrian. 2013. Boolean Minimization in Social Science Research: A Review of Current Software for Qualitative Comparative Analysis (QCA). *Social Science Computer Review*, *31*(4), 505–521.