

Do subnational performance assessments lead to improved governance? Evidence from a field experiment in Vietnam

Using SPAs to improve governance outcomes

117

Edmund Malesky, Tuan-Ngoc Phan and Anh Quoc Le
Department of Political Science, Duke University, Durham, North Carolina, USA

Received 7 February 2022
Revised 26 April 2022
Accepted 6 July 2022

Abstract

Purpose – Single-party regimes increasingly use Subnational Performance Assessments (SPAs) – rankings of provinces and districts – to improve governance outcomes. SPAs assemble and publicize information on local government performance to facilitate monitoring and generate competition among officials. However, the evidence are sparse on their effects in this context. The authors argue that built-in incentive structures in centralized single-party regimes distort the positive impact of SPAs.

Design/methodology/approach – The staggered rollout of the Vietnam Provincial Governance and Public Administration Performance Index (PAPI) created a natural experiment. Due to 2010 budget constraints, the first iteration of the PAPI survey covered only 30 of Vietnam's 63 provinces before covering all in 2011. The PAPI team used matching procedures to identify a statistical twin for each province before randomly selecting one from each pair. The authors use randomization inference to compare the outcomes of these control and treatment groups in 2011.

Findings – Exposure to PAPI helped improve almost all aspects of governance; however, significant evidence of prioritization bias exist. The positive effects only persisted for the dimension of administrative procedures, which was the one area of governance that was prioritized by the central government at the time. Other dimensions only registered short-term effects.

Originality/value – Our study provides an examination of the impact of SPAs in a single-party regime context. In addition, the authors leverage the natural experiment to identify information effects causally. The authors also look past short-term effects to compare outcomes for five years after the treatment occurred.

Keywords Vietnam, Transparency, Decentralization, Authoritarianism

Paper type Research paper

1. Introduction

Authoritarian regimes worldwide are increasingly using Subnational Performance Assessments (SPAs), rankings of the provinces and districts within a country, to improve local governance outcomes [1], [2]. The driving force behind this recent trend has been the combination of a need to improve governance and a critical monitoring problem. Despite their authoritarian nature, these regimes must deliver sufficient economic success and governance quality to maintain legitimacy (Gandhi, 2004; Magaloni, 2006; Przeworski & Limongi, 1993; Whiting, 2017). As in most principal-agent relationships, however, despite having the power to punish low-performing subordinates, a top leadership has an informational deficit with regard to lower-level bureaucrats, which is a problem made worse by their nondemocratic tendencies (Kung & Chen, 2011; Wintrobe, 2000; Yang, Xu, & Tao, 2014). Regardless of the specific regime type, information drives the causal logic behind SPAs by empowering voters or central leaders with the tools to monitor how well subnational officials have achieved



© Edmund Malesky, Tuan-Ngoc Phan and Anh Quoc Le. Published in *Fulbright Review of Economics and Policy*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Fulbright Review of Economics
and Policy
Vol. 2 No. 2, 2022
pp. 117-135
Emerald Publishing Limited
e-ISSN: 2635-0181
p-ISSN: 2635-0173
DOI 10.1108/FREP-02-2022-0006

governance goals. By filling this information gap, SPAs enable principals to effectively wield their power and incentivize agents to perform.

The use of SPAs in authoritarian regimes is consistent with claims that political selection in certain authoritarian regimes is meritocratic, built on a yardstick competition, and can deliver good governance outcomes (D. Bell, 2015; D.A. Bell, 2015; Bell, 2015; Jia, Kudamatsu, & Seim, 2015; Landry, Lü, & Duan, 2017). In this paper, we argue that built-in incentive structures in centralized authoritarian regimes can distort the positive impact of SPAs. In particular, since local leaders are more accountable to the central party hierarchy than to regular citizens in such systems, their reform efforts will be disproportionately directed toward areas prioritized by the top leaders rather than those valued by society at large.

We evaluate the effects of the Vietnam Provincial Governance and Public Administration Performance Index (PAPI) in Vietnam, an internationally lauded SPA funded by the United Nations Development Programme (UNDP). PAPI was profiled by Princeton University's Innovations for Successful Societies program as an example of a successful social audit of governance reform and by the World Bank as a model for active citizenship (Oxfam, 2012; Jackson, 2014). Other projects in Thailand and Mexico have also adopted the PAPI methodology (VOV, 2013; Jukic, 2015). Existing research testifies to PAPI's effectiveness in spurring improvements in governance (Giang, Nguyen, & Tran, 2017).

In single-party regimes like that of Vietnam, we argue that SPAs improve governance outcomes by serving as a monitoring tool for the central government. In the case of the PAPI index, promotion to higher office is the primary material motivation driving the decisions of local leaders. Accurate information about the performance of subnational units allows the central government to reward local officials accordingly. It also allows for reactivity as provincial leaders learn their relative performance as benchmarked against peers and make policy changes in response to the new information. Therefore, a regularized, publicized index can motivate actions by local leaders and lead to improvements in the measured outcomes.

However, classic works on principal-agent relationships have shown that when an agent's task is multidimensional, incentives can lead to suboptimal outcomes even under perfect information, such as when good, objective output measures are available (Holmstrom & Milgrom, 1991). When there are multiple tasks, not only do information and incentives motivate hard work, they also affect how agents allocate their efforts among their various duties. Specifically, we theorize and empirically confirm that the importance of material consequences generates a *prioritization bias*. Subnational officials dedicate more attention, energy and resources to what they believe those using the index want. In our case, provincial leaders focused on improvements to the quality of Administrative Procedures, which were a priority of the central government at the time of our experiment. Some provinces had a head start having been randomly chosen to participate in the PAPI 2010 report and thus received early feedback, while the rest entered in 2011. The former continued to achieve higher scores for the administrative procedures dimension long after the initial advantage was taken away. In fact, we could still detect the SPA's impact up to five years after PAPI had been implemented across the entire country. By contrast, for governance objectives that did not receive commensurate attention from Vietnamese central officials, the treatment effect, if any, was short-lived, with the new entrants quickly catching up.

Research on prioritization bias in the public administration literature has focused predominantly on Western, democratic regimes. Scholars show that performance measurement has a small, positive effect on governance quality (Gerrish, 2016), and that the effect is contingent upon supportive interest groups or institutions (Kroll, 2017). On the other hand, in both democratic and nondemocratic settings, studies have documented many examples of officials gaming the system and achieving governance goals on paper but not in spirit (Jacob & Levitt, 2003; Gao, 2009; Guo, 2009; Heinrich & Marschke, 2010; Chan & Gao, 2012; Wallace, 2014; Li, 2015). Existing research has also provided evidence that fixating on

numerical targets can lead to unintended outcomes, from worsened pollution (Shen, 2018) to political radicalism and the loss of human life (Kung & Chen, 2011). In this paper, we show that even without deception and malevolent intentions, when an authoritarian regime uses SPAs to incentivize performance, it suffers from exactly the same problems that afflict bureaucracies everywhere.

2. The theoretical goals of SPAs

2.1 *Yardstick competition in centralized and authoritarian regimes and SPAs as the yardstick*

Centralized and authoritarian systems have enthusiastically adopted the use of SPAs. Doing Business reports the subnational indices of China, Russia and Egypt, while The Asia Foundation has created economic governance indices for Vietnam, Malaysia, Cambodia and Bangladesh and is preparing to roll out another one for Myanmar (The Asia Foundation, 2011; World Bank, 2017). To understand how SPAs affect the behaviors of subnational units, it is helpful to return to the rich political economy literature on decentralization in democracies. Economists have argued that greater local authority in decision-making improves the efficiency of public service delivery because government output can be provided in small units and tailored directly to local tastes (Oates, 1972; Besley & Coate, 2003). In addition, decentralization creates competition for capital and labor that leads to improved governance outcomes, brings decision-making closer to citizens and local businesses, and limits the role of central government interventions in economic policies (Tiebout, 1956; Rondinelli, McCullough, & Johnson, 1989; Inman & Rubinfeld, 1997). Subnational competition is also thought to encourage policy innovation, which is a key feature of market-preserving federalist arguments (Weingast, 1995).

In more centralized systems, the theoretical role for SPAs shifts slightly. While the motivation for facilitating competition for labor and capital remains, the goal of empowering local voters is less pronounced as they do not have a direct means to sanction local leaders. Advocacy possibilities remain available but must appeal to the incentives of central or local leaders. Instead of operating through downward accountability to constituents, the monitoring and reactivity mechanism in centralized and authoritarian systems leverages upward accountability to central authorities. In areas where their goals and those of the index architects align, central authorities can use SPAs to monitor and then reward local officials who perform better on objectives they prize. Our focus on upward accountability thus provides a different perspective from that of other studies on performance assessments whose mechanisms are bottom-up pressure from voters and societal groups (Courty & Marschke, 2004; Kroll, 2017).

The role of SPAs in augmenting yardstick competition between localities may be particularly effective in single-party regimes, such as in Vietnam. In the blossoming literature on authoritarian institutions, one explanation for the better economic performance and longevity of single-party regimes over other types has been the role of party organization in promotion and advancement (Svolik, 2012). Low-level members are forced to invest in party service and meet party goals, which pay off in terms of party promotion but have little value elsewhere. Party goals involve outlined and complicated systems that are put into place to determine whether officials have achieved these goals. This interjurisdictional yardstick competition promotes party objectives and breeds party loyalty as long as sufficient space is made available for advancement (Lazarev, 2005; Stiglitz, 1989). While performance is rewarded in many regimes, the critical difference between single-party states and other authoritarian systems is the formal promotion system, including review institutions, such as the Organization Department of the Chinese Communist Party, and clear promotional ladders from functionary to elite levels. Such codification makes these performance criteria more credible and enticing (Magaloni, 2008; Reuter & Turovsky, 2014).

In China, the in-depth literature on regionally decentralized authoritarianism has included the study of the effectiveness of cadre evaluation criteria and demonstrated the strong correlations between local economic performance and promotion. However, even in a high-capacity state such as China, doubts have been raised about the objectivity and quality of performance criteria (Guo, 2009; Wong, 2016).

In settings where capacity is low or there are doubts about the independence of government research institutes, the SPAs generated by external experts can provide the objective yardstick that central leaders need to make personnel decisions. For instance, in our setting of Vietnam, the Provincial Competitiveness Index (PCI), a ranking of the business environments of provincial governments that was funded by US-AID and created by the Vietnamese Chamber of Commerce and Industry was included in Decree 19 as an official indicator of economic reform progress (Government of the Socialist Republic of Vietnam, 2014). Anecdotally, the index has created incentives for local leaders to up their game. All 63 provinces of Vietnam have released PCI action plans (Cong-Huong, 2020). As recently as June 2021, the Hanoi government issued a directive calling for greater effort from its bureaucrats to improve the capital's PCI scores. The directive drew attention to two low-performing PCI subindices and proposed specific solutions (Hoai-Thu, 2021). Toward the other end of the country, the Chairman of Dong Thap People's Committee, a province in the Mekong Delta, considered PCI subindices as "arteries of the economy" (Huu-Nghia, 2021). Thus, even in nondemocratic regimes, SPAs can enhance the governance outcomes of subnational units when there is some form of quasi-meritocratic promotion. From this insight, we generate our first hypothesis as follows:

- H1. Subnational officials subjected to SPA rankings are more likely to improve the governance performance of their localities as compared to subnational officials left out of the rankings.

2.2 Prioritization bias in the allocation of efforts by subnational officials

The above discussion indicates the potential for SPAs to help improve performance when they are aligned with the interests of the central leaders of single-party regimes. However, it also highlights the limitations of SPAs in such settings. When SPAs rely on the mechanism of upward accountability to central benefactors in order to generate change, the expressed goals of the central leaders will play the largest role in the prioritization by subnational officials. This issue is well documented in the management literature on devising appropriate incentive structures for employees (Koch & Peyrache, 2011) and public administration literature on performance targets (Boyne & Chen, 2007). While determining how to budget scarce time and economic resources, subnational officials will likely concentrate their reform interventions in the areas that they know are the most likely to be rewarded. Subindices and indicators deemed less important or at odds with the interests of the central authorities will receive less attention. Therefore, the lasting impact of SPAs is likely to be observed only in aspects of performance that are important to the central government. We call this phenomenon prioritization bias, which motivates our second hypothesis.

- H2. Subnational officials subjected to SPA rankings are more likely to achieve lasting improvements in the areas of governance prioritized by central authorities.

3. Field experiment setting

To test causally whether SPAs improve governance, we leverage the staggered, randomized rollout of a provincial governance index in Vietnam to establish genuine treatment and control groups. The index, called the Vietnam Provincial Governance and Public

Administration Index (PAPI), was created in 2010 by the UNDP and the Center for Research and Training of the Vietnam Fatherland Front (VFF) to measure local governance along six dimensions: (1) participation at local levels, (2) transparency, (3) vertical accountability, (4) control of corruption, (5) public administrative procedures and (6) public service delivery (Centre for Community Support and Development Studies [CECODES] *et al.*, 2011) [3].

While the UNDP is a well-known multilateral developmental organization, not all readers may be familiar with the VFF. The VFF is the umbrella organization for Vietnam's mass organizations, including the Labor Confederation, Women's Union, Peasant Union, Youth Union and Communist Party, among others (Jeong, 1997; Thayer, 2009). In single-party systems based on Leninism, organizations like the VFF play the dual role of (1) downward communication from the central elites toward its member organizations to "create consensus" and (2) upward reporting of its members' views and needs to central policy-makers to "tighten intimacy between people and government." The leaders of the VFF like to cite their listing in the 1992 Constitution as the "political base of the people's power" and suggest that they are also meant to monitor the activities of local governments on behalf of their leadership [4]. The VFF is a silent partner in the research and construction of PAPI. To avoid biasing interviews, its staff does not accompany interviewers and its name is not mentioned in the introductory lesson to the survey. It also plays no role in the calculation of scores or in the authorship of the report. The VFF's key role is using its network of local bodies to obtain permission for conducting the survey throughout the country [5]. Despite its limited research role, the VFF's name does appear on the cover of PAPI reports. Its authority is an important reason why local officials associate the index with the demands of the central leadership.

It is important to note that, while PAPI creates an aggregate index, the annual report relegates it to an appendix. The formal presentation in Vietnam prefers a dashboard approach where rankings on the six dimensions are presented separately in a star chart. Staying true to the goals of PAPI, we also study the individual dimensions rather than the aggregate score, which combines topics that are too theoretically distinct to be analytically meaningful.

In 2010, the PAPI report was not fully inclusive, which presents us with an important opportunity that we exploit in our research design. Due to budget constraints in 2010, the UNDP decided to cover only 30 of Vietnam's 63 provinces with a goal of demonstrating success before seeking additional funding (CECODES *et al.*, 2011). To ensure the representativeness of the initial provinces, the PAPI team used matching procedures to identify a statistical twin for each province in the country. Then, the researchers randomly selected one province from each pair and included it in the PAPI survey while keeping the other in the control group. In total, PAPI surveyed 5,560 citizens from 30 provinces in 2010. In contrast, from 2011 onward, PAPI has covered all 63 provinces with 13,642 responses.

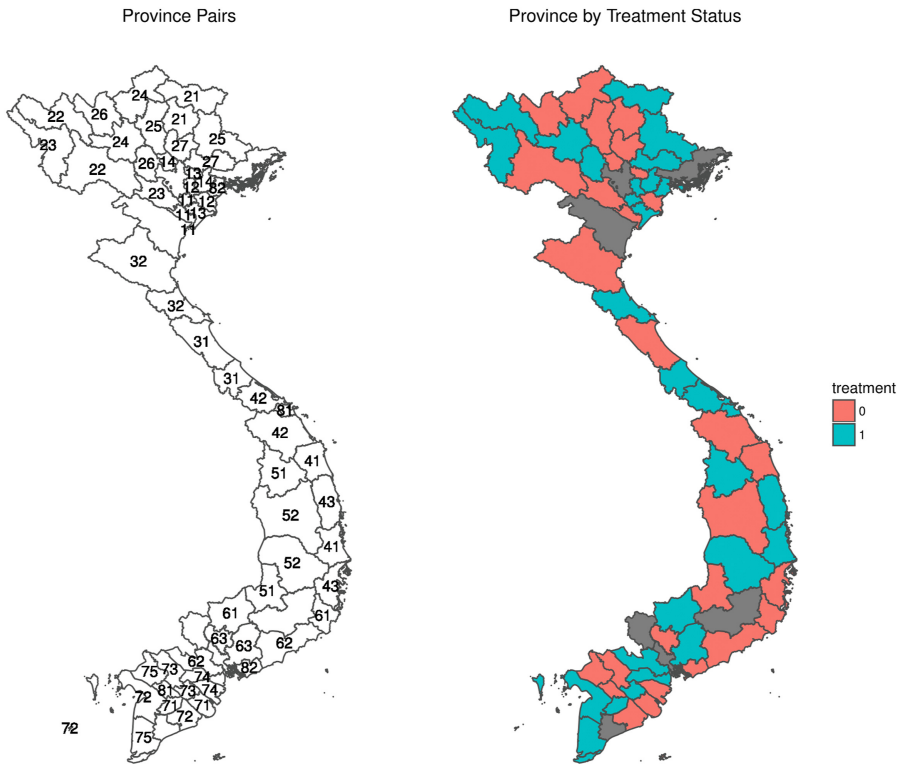
3.1 Treatment

Our treatment is "being included in the PAPI 2010 survey." The numbers in the left panel of Figure 1 show the pairs to which the provinces belong. The colors in the right panel of Figure 1 show the treatment statuses of the provinces assigned by the researchers after randomizing within the pairs. Several provinces were deliberately selected into or excluded from PAPI 2010. For example, for policy relevance, the researchers deliberately ensured the inclusion of Hanoi and Ho Chi Minh City, the two biggest cities in Vietnam. On the other hand, Quang Ninh, Thanh Hoa, Lam Dong, Tay Ninh and Bac Lieu were excluded due to the lack of a good match. We remove these nonrandomly assigned provinces from all analyses, resulting in a sample of 56 provinces, with 28 in the control and the other 28 in the treatment group. To the best of our knowledge, this is the first time that an SPA was designed in such a way that a legitimate control group is available for comparison.

Table 1 shows the covariate balance between the control and the treatment groups for a variety of structural and institutional variables. Since governance and development are highly correlated, we check the balance of various proxies for the level of development, including GDP per capita, percentage of asphalt roads, telephones per capita and other data. In addition, because past governance is the best predictor of future governance, we check the balance of pretreatment governance quality. Even though PAPI data are not available before its pilot in 2010, we can use the annual PCI, a measure of a province's business environments. Table 1 shows that all of these covariates are balanced as expected from the randomized design.

3.2 PAPI's Dimension 5: administrative procedures quality

We theorize that SPAs provide the central government with an informational tool to monitor provincial governments. Therefore, we expect to see persistent effects only in governance areas that (1) the central government prioritizes and (2) are measured directly in the SPAs. One area that satisfies both conditions is PAPI's Dimension 5, administrative procedures. Indeed, the interest in administrative reform has been a running theme in Vietnam's modernization. As the 1986 economic reform (*doi moi*) began, the government was acutely aware that its rapidly transforming economy needed a modern, efficient public



Note(s): *(Left) The paired numbers indicate which two provinces belong to the same pair. (Right) The colors indicate the treatment statuses of the provinces. Uncolored provinces were selected into the treatment or control groups deliberately. We remove these provinces from all analyses because their treatment statuses are not randomly assigned

Figure 1.
Treatment assignment

Table 1.
Balance table

Variable	Control mean	Treatment mean	Diff	<i>p</i> -value
Agricultural share (%)	33.31	34.44	1.13	0.77
Asphalt roads (%)	73.81	74.81	0.38	0.95
Distance to Hanoi or HCMC (km)	264.00	282.79	18.79	0.76
GDP per capita (million VND)	19.50	14.60	-4.90	0.28
Population (thousands)	1237.31	1359.29	121.98	0.66
Secondary school graduates (%)	74.84	78.89	4.05	0.27
Telephones per capita	0.21	0.21	0.00	0.99
Business environment (measured by PCI)	52.91	51.50	-1.41	0.23

Note(s): *As expected from the randomized design, we have excellent balance across covariates. The size of the differences is substantially small and statistically insignificant, with all *p*-values larger than 0.2. 1 million VND = 50 USD

administration. The result was the Master Program on Public Administration Reform (PAR) for the period 2001–2010, with four key reform areas: (1) institutional reform, (2) organizational structures, (3) civil service reform and (4) public finance (Painter, 2005). Throughout its conception and development from 1991 to 2001, PAR received continuous endorsement from the Vietnam Communist Party (VCP) as articulated in the party directives (Seventh and Eighth Party Congresses) (UNDP, 2001).

Halfway through the period from 2001 to 2010, PAR received a second wind from the support of Prime Minister (PM) Nguyen Tan Dung. Throughout his career, Dung had risen as a proactive politician, pushing for the United States–Vietnam Bilateral Trade Agreement and Vietnam’s World Trade Organization accession over a divided VCP. As a newly chosen PM, he adopted PAR as his next signature project. Dung formed a special task force to coordinate Project 30 (*De an 30*), an initiative so-called because of its aim to review all administrative procedures and to cut at least 30% of compliance costs for businesses and citizens (Schwarz, 2010). Project 30 was a high-profile initiative with much at stake. PM Dung assigned the project to his protégé Nguyen Xuan Phuc, who was then the head of the Government Office before going on to serve as PM himself, succeeding Dung [6]. In addition, the timing of Project 30 is especially telling regarding its political implications for PM Dung. After three years of preparation, Project 30’s five-month implementation phase ended in November 2010, a mere two months before the 11th Party Congress (January 19, 2011), which would determine Dung’s second term. It was clear that Dung was personally invested in the project. As the deputy manager of Project 30 special task force put it, “There has never been any administrative reform that was as closely dictated by the PM as Project 30” (Minh-Hang, 2009).

Not only was the central government invested in administrative reform, we suggest that it was especially interested in the information that PAPI 2010 provided. Indeed, while administrative reform had been a priority, it was very difficult for the central government to accurately assess the performance of the provinces. According to Acuna-Alfaro, the UNDP’s Governance Officer responsible for PAPI, there was a clear “. . . lack of a monitoring and evaluation system not only for the overall [PAR] master plan but also for overall governance dynamics in Vietnam. The government had recognized the need for a monitoring system for the PAR program but had failed to set it in motion” (Jackson, 2014). Much of the information that did exist came from the provinces themselves, who were eager to paint a flattering portrait of their own performances. According to the Deputy Director of the Department of PAR, “the Government [. . .] had no proper tools to assess in a scientific, objective, and quantifiable manner whatever outcomes that have been achieved from the PAR implementation [. . .] The work of assessment remained qualitative, subjective and yet lacked a participative approach from the civil society” (Pham, 2010). Indeed, PAPI was

conceived partly to address this very need of assessing public administrative performance from the citizens' eyes (CECODES *et al.*, 2010).

Crucially, since Project 30's implementation and PAPI 2010 both took place during the latter half of 2010, PAPI was able to provide the central government with the timely information it both wanted and needed. In addition, the narrow focus of Project 30 on citizen-facing administrative procedures aligned perfectly with PAPI's Dimension 5. Knowing that the central government could accurately detect and punish poor performers, the treated provinces were incentivized to do well. Therefore, we expect their performance in administrative procedures to have improved.

3.3 Administrative reform as the most prioritized governance area

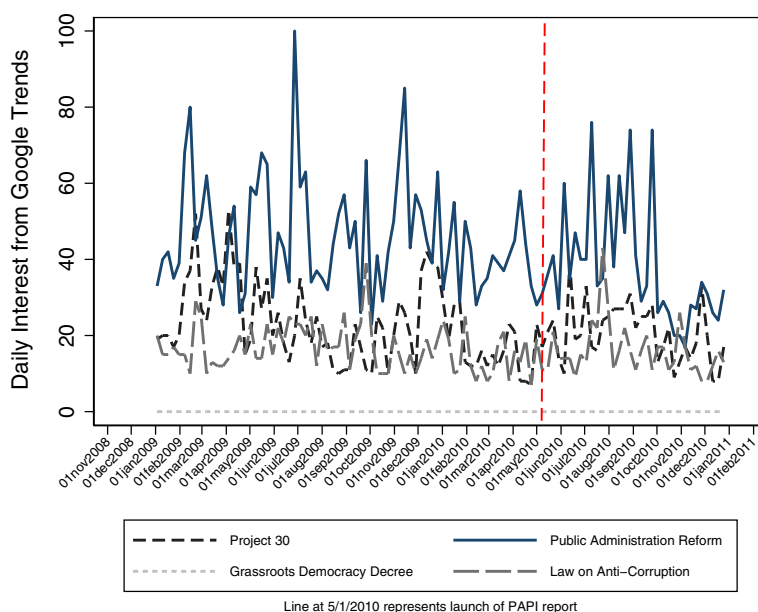
In addition to measuring the quality of administrative procedures, PAPI surveys citizens' experiences with political participation, transparency, vertical accountability, corruption and public service delivery. However, at the time of the 2010 treatment, administrative reform in general and Project 30 in particular were the central government's foremost priorities. Since Project 30 was PM Dung's personal project, it was able to bypass the regular legislative channels and had the full endorsement of the executive branch.

Although it is difficult to definitively know the priorities of the Vietnamese leadership, Schuler (2014) has demonstrated a strong correlation between the priority topics in the Vietnamese National Assembly and the interest of the Vietnamese public. He documents this association using text analysis of Vietnamese newspapers and Google Trends data. Following this approach, in Figure 2, we compare interest in Project 30 and PAR to two other legal documents that informed the construction of the PAPI subindices [7]. The y-axis records the amount of daily interest for the search term using the Google Trends algorithm, which has a range between 0 and 100 with 100 corresponding to the most searched term on a particular day. It is immediately clear that PAR, marked by the thick solid lines, dominates the other documents. Of the four legislative programs, it is the most searched term every single day. Project 30 is the second most searched document over the time period, followed by the Anti-Corruption Law. Certainly, at the time of the PAPI launch in May 2010, PAR was prominently on the minds of Vietnamese Internet users.

One may argue that the difference between administrative procedures and other governance dimensions is not so much how prioritized they are, but how easy it is to change them. In other words, we may see changes in administrative procedure quality because it is easy to remove existing paperwork. While we find this explanation plausible, we argue that it is difficult to determine *a priori* whether change is hard or easy. Indeed, removing paperwork may take away important sources of rents from provincial bureaucrats, generating insurmountable resistance. Lacking a clear theoretical guideline, we decide not to engage in *post hoc* data mining to show "proof" that some governance areas are easier to change than others.

3.4 Operationalization of administrative procedures quality

To operationalize this principal dependent variable, we use citizen responses to the PAPI survey regarding experiences with public administration procedures. Since 2011, PAPI has conducted the survey in all 63 provinces, allowing us to compare citizen assessments in control and treatment provinces. We do this using the PAPI subindex on administrative procedures, which aggregates citizen experiences with the four most common procedures for Vietnamese citizens: (1) public certification, which is Vietnam's equivalent of public notary services; (2) construction permits granted to civil construction projects (e.g. building, expanding or remodeling houses in anything more than a basic way); (3) application procedures for the renewal and transfer of land use right certificates for citizens, granting them a quasi-property right that is exchangeable and mortgageable and (4) administrative procedures for personal papers (e.g. birth certificates,



Note(s): *The priority of the Vietnamese central leadership among public administration reform, anti-corruption, and grassroots democracy as approximated by public interest. The y-axis displays the level of public interest, ranging from 0 to 100, with 100 assigned to the most searched term of the day. “Public administration reform” and “Project 30” almost always rank as the top two searched terms

Figure 2.
Google Trends

marriage certificates, death notifications, residency registrations and welfare subsidies). These procedures are selected from the list of administrative procedures that local governments are delegated to process for citizens.

For each procedure, citizens respond to eight yes/no questions regarding their experiences with local authorities in receiving documents. These questions, which can be seen in [Table 2](#), cover: (1) the clarity of the application procedures, (2) publicity of application fees, (3) competence of civil servants, (4) behavior of civil servants, (5) reasonable paperwork load, (6) notification of deadlines, (7) receipt of results within the set deadline and (8) overall service satisfaction.

This battery of questions is an appropriate operationalization of a prioritized governance area for three reasons. First, theoretically, the questions directly map onto the government’s own criteria for reform under the PAR program. Second, methodologically, the questions reflect PAPI’s focus on citizen experiences, not perceptions. PAPI has filtering questions to keep only the answers of citizens who went through the administrative procedure. The choice of dichotomous answers limits perception and anchoring biases caused by Likert scales. In addition, citizens are only asked about aspects of the procedure that they have directly observed and can understand. Third, the question regarding whether informal fees (i.e. bribes) were paid is used in the PAPI measurement of corruption (Dimension 4) and not in the analysis of administrative procedures.

PAPI aggregates the answers to all 32 of these questions (four procedures with eight questions each) into a ten-point scale that they present at the province level. In 2015 for instance, the median Dimension 5 score was 6.83, with Quang Ngai receiving the lowest score of 5.9 and

D503. Below are some assessments about the certification process at the local government. From your own experience, do you think these assessments are correct or incorrect?

	Correct	Incorrect
a. I received specific guidances about the process	1	0
b. Fees were posted publicly	1	0
c. The staff was competent	1	0
d. The staff was polite	1	0
e. The paperwork was simple	1	0
f. There was no need to pay informal fees	1	0
g. I was given a date to pick up the results	1	0
h. I received the results on time as promised	1	0
i. I was overall happy with the service	1	0

Table 2.
Questions used by PAPI to measure citizen experiences with administrative procedures

Note(s): *To make sure that our results truly reflect changes in experience and not perception, we also construct our own index using only responses for the objective questions: a, b, g, and h (marked red)

Bac Ninh receiving the highest of 7.51 (CECODES *et al.*, 2016, p. 70). In the construction process, however, PAPI actually calculates Dimension 5 scores for each individual respondent based on their survey questions. This allows them to aggregate to different levels of government (e.g. district, commune, village) and also assess through regression analysis whether different types of citizens (i.e. females, ethnic minorities) experience governance differently.

4. Empirical model

In our experimental design, the treatment is randomized at the provincial level while the outcome is measured at the individual level with survey weights. Such a design affects the standard errors of our estimates in two ways. First, the correlation between individuals in the same province enlarges the standard error. Second, the complex survey weights in PAPI further complicate how to calculate the standard errors.

While provincial-level randomization and complex survey sampling are straightforward to deal with separately, when they are both present, it is unclear how to manage their countervailing effects on the efficiency of estimates [8]. This difficulty is common among field experiments, where the treatment is often randomized at the cluster level (e.g. village, school) and the outcome is measured by a survey at the individual level (e.g. villagers, students).

To resolve this issue, we use an approach based on randomization inference (RI) to take into account both the survey sampling weights and provincial-level treatment. RI takes advantage of the fact that the researchers who conducted the experiment know the precise procedures by which the units are assigned to the treatment and control groups. Therefore, we can simulate all other possible random assignments, measure the treatment effects in those scenarios and place the actual experimental result against the backdrop of all other hypothetical results. RI is particularly suitable for analyzing experiments in cases of complex randomization procedures and the clustering of observations, such as in this setting (Gerber & Green, 2012). Moreover, the

nonparametric nature of RI means that the results are less dependent on the researchers' model specification choices (Imbens & Rubin, 2015).

Specifically, we follow a three-step procedure:

- (1) First, we randomly reassign the individual respondents to treatment and control statuses 1,000 times in accordance with PAPI's pair randomization scheme.
- (2) Then, using the original outcomes and covariates, we recalculate the *hypothetical* treatment effect in each of these 1,000 permutations or scenarios. This provides us with a distribution of re-randomized estimates.
- (3) Finally, we compare our original *observed* effect size to this distribution. If the actual experimental estimates lie in the extremes of this range, it is highly unlikely that the treatment effect is zero and that our results are simply due to chance. For example, if the original effect size is higher than 399 out of every 400 re-randomized results, the effect is considered statistically significant at the 0.005 level.

The use of RI offers three benefits. First, it solves the thorny issue of adjusting inference in the presence of both intra-province correlation and survey weights. Second, the RI test can simulate the experiment in a manner that perfectly mirrors how researchers randomized the treatment in reality. In this case, our RI test follows PAPI's pair randomization scheme, which offers important advantages over complete randomization for studies with a small sample size like ours. It ensures that, within a pair of similar provinces, we will always assign one province to the control and one to the treatment group. Therefore, we guarantee that the composition of the control and treatment groups will be similar, giving us more confidence that the difference in outcomes is not due to a bad draw assigning all high-performing provinces to the treatment group. A standard regression approach would fail to take advantage of this pair randomization structure.

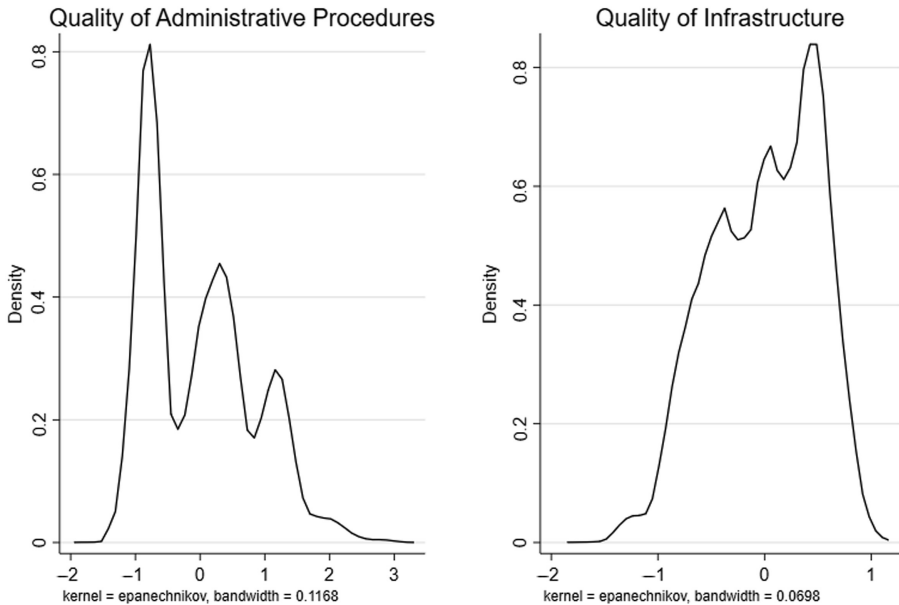
Third, the nonparametric nature of the RI test is desirable because our data do not satisfy the assumption of a normally distributed error variance of an ordinary least square (OLS) approach. Our outcome indices are constructed to be bounded between 1 and 10, causing many observations to cluster near the bounds, which will likely lead to non-normal error distributions. Figure 3 confirms this concern. We plot the distribution of the residuals from an OLS regression of administrative procedure quality and infrastructure quality against the treatment, showing a skewed distribution of the residuals. In this situation, it is more defensible to use a nonparametric approach such as RI. For the results of the RI test to be valid, we do not need any modeling assumptions except that the treatment is randomized, which is satisfied by the design.

5. Empirical results

5.1 Overall impact

Table 3 demonstrates the effect of being included in the PAPI 2010 survey on outcomes in 2012. Individuals in the treated provinces reported PAPI scores that were 0.575 higher on average than those in provinces that were not included in 2010. Although not a huge effect, an increase of this magnitude is equivalent to more than one-fourth of the 2012 standard deviation in PAPI scores across provinces (1.99). Moreover, average PAPI scores nationwide only increased by two points, from 35.37 to 37.65, from 2012 to 2020. The immediate impact of inclusion in PAPI 2010 on 2012 scores is quite impressive when put against this background of overall glacial improvement.

Digging deeper, we found evidence that this SPA helped improve almost all aspects of governance measured. Scores in the corruption dimension increased the most with 0.191 followed by administrative procedures (0.107) and accountability (0.103). These are



Note(s): *The figure illustrates the skewed distributions of the error terms in the OLS regression. We regress the administrative procedures quality (left) and the infrastructure quality (right) on the treatment and then plot the density of the residuals. The plots show that the residuals are not normal, violating the normality assumption OLS regression. In this situation, a non-parametric approach such as RI is more appropriate

Figure 3.
Residual plots

substantial changes given that the average scores in these three dimensions are 5.84, 6.87 and 5.58, respectively. Citizen evaluation of public services also improved (0.068). The fact that control of corruption is the most improved area comes as a surprise, considering the intense focus on PARs at the time. However, in the next section, we show that sustained progress is only observed in administrative procedures. The other PAPI dimensions only registered short-term effects.

5.2 Persistence and prioritization bias

Table 4 demonstrates that the treatment effect is short-lived for most PAPI dimensions. The impact on accountability, control of corruption and public services dissipated immediately in 2013. The short-term spike in these scores seems to have come from an audience effect on provincial officials rather than as a result of an indigenous reform effort.

The positive impact only persisted in one particular PAPI dimension, administrative procedures. Being included in 2010 led to a substantially higher score in this dimension in 2011 (0.135). This head start advantage did diminish over time, decreasing to 0.107 and 0.082 in 2012 and 2013, respectively. However, there was still a statistically significant difference between treated and untreated provinces five years later (0.065) [9].

This gap is particularly striking given how quickly any differences in the other dimensions dropped to zero after 2012. In 2013, while the gap between treatment and control groups narrowed to 0.01, 0.009, 0.037, 0.019 and -0.023 for participation, accountability, transparency, corruption and public services, respectively, the advantage remained at 0.082 for administrative procedures. Given our very demanding threshold for statistical

Dependent variables	Weighted PAPI	Unweighted PAPI	D1: Participation	D2: Accountability
<i>Average treatment effect</i>	0.575*	0.469*	-0.005	0.103*
Standard <i>p</i> -value	(0.02)	(0.102)	(0.942)	(0.225)
Holm <i>p</i> -value	(0.007)	(0.01)	(0.025)	(0.017)
Randomization inference <i>p</i> -value	(0.000)	(0.000)	(0.878)	(0.000)
Mean in control	38.576	36.993	5.391	5.905
Standard deviation in control	4.952	4.956	1.316	1.506
Observations	11,086	11,086	11,086	11,086
<i>R</i> ²	0.002	0.003	0.004	0.003

Dependent variables	D3: Transparency	D4: Corruption	D5: Administrative procedures	D6: Public services
<i>Average treatment effect</i>	0.005	0.191*	0.107*	0.068*
Standard <i>p</i> -value	(0.957)	(0.027)	(0.01)	(0.172)
Holm <i>p</i> -value	(0.05)	(0.008)	(0.006)	(0.013)
Randomization inference <i>p</i> -value	(0.83)	(0.000)	(0.000)	(0.000)
Mean in control	5.806	5.840	6.990	7.062
Standard deviation in control	1.309	2.011	0.858	0.730
Observations	11,086	11,086	11,086	11,086
<i>R</i> ²	0.002	0.001	0.004	0.003

Note(s): *Significant at the 0.005 level using RI, meaning that in 1,000 permutations, five or fewer had greater average treatment effects (ATEs) than the observed result. Standard errors clustered at the provincial level with fixed effects for the matched pairs used in the pair randomization sequences

Table 3. The effect of inclusion in PAPI 2010 on 2012 scores

significance ($p = 0.005$), it is highly unlikely that the persistent effects of the treatment on public administrative procedures occurred by accident. This finding is consistent with our hypothesis that PAPI uniquely spurs provincial politicians to achieve substantive changes in areas championed by their superiors in Hanoi.

Next, we dig deeper into the impact of the treatment on administrative procedures to explore the inner workings. This PAPI dimension comprises four subindices, including certification, construction permits, land procedures and other procedures [10]. According to Table 5, the persistent effect observed above is driven by improvements in construction permits and land titling/exchange procedures. The citizens in the provinces that were included in PAPI 2010 continued to report better experiences acquiring construction permits and land titles five and four years later, respectively.

The fact that progress concentrating in these two areas is not surprising. While public certification mainly involves relatively straightforward procedures, the issuance of construction permits and land titles involves more complicated processes with more room for improvement. It is therefore understandable that provinces can achieve higher scores by focusing their efforts on subindices with low starting points rather than trying to upgrade areas already close to their upper limits. In 2014, on a scale of 1 to 8, the national average quality score for public certification was 7.30. In contrast, the corresponding figures for construction permits and land procedures are 6.66 and 5.04, respectively.

While the treatment effect on these subindices is substantively meaningful, they are subject to a noisy measurement. Even though these are the four most common administrative procedures with which Vietnamese citizens are familiar, it is rare for an individual to experience all of them in any particular year. A respondent's answers to these four subindices

Dependent variables	Weighted PAPI	Unweighted PAPI	D1: Participation	D2: Accountability
ATE in 2011 (<i>n</i> = 10,949)	0.194 (0.035)	0.281 (0.02)	0.002 (0.933)	-0.022 (0.45)
ATE in 2012 (<i>n</i> = 11,086)	0.575* (0.000)	0.469* (0.000)	-0.005 (0.878)	0.103* (0.000)
ATE in 2013 (<i>n</i> = 11,122)	0.002 (0.978)	0.134 (0.136)	0.010 (0.682)	0.009 (0.746)
ATE in 2014 (<i>n</i> = 10,799)	0.111 (0.336)	0.077 (0.421)	-0.029 (0.280)	0.079 (0.06)
ATE in 2015 (<i>n</i> = 10,790)	0.083 (0.498)	0.094 (0.353)	-0.042 (0.122)	0.092* (0.000)
Mean of control in 2011	36.38	36.54	5.05	5.70
SD of control in 2011	5.93	5.01	1.41	1.51

Dependent variables	D3: Transparency	D4: Corruption	D5: Administrative procedures	D6: Public services
ATE in 2011 (<i>n</i> = 10,949)	0.103* (0.000)	0.006 (0.88)	0.135* (0.000)	0.057* (0.000)
ATE in 2012 (<i>n</i> = 11,086)	0.005 (0.83)	0.191* (0.000)	0.107* (0.000)	0.068* (0.000)
ATE in 2013 (<i>n</i> = 11,122)	0.037 (0.135)	0.019 (0.599)	0.082* (0.000)	-0.023 (0.087)
ATE in 2014 (<i>n</i> = 10,799)	0.042 (0.098)	-0.018 (0.640)	0.039 (0.015)	-0.035 (0.013)
ATE in 2015 (<i>n</i> = 10,790)	0.076 (0.010)	-0.047 (0.243)	0.065* (0.000)	-0.05 (0.001)
Mean of control in 2011	5.68	5.97	6.87	7.27
SD of control in 2011	1.63	2.04	0.88	0.70

Table 4.
The effect of inclusion in PAPI 2010 on historical PAPI scores

Note(s): *Significant at the 0.005 level using RI, meaning that in 1,000 permutations, five or fewer had greater ATEs than the observed result. Coefficients and *p*-values (in parentheses) are generated by RI clustered at the provincial level. We include fixed effects for the matched pairs used in the pair randomization sequences

Dependent variables Subdimensions of admin. procedures	Certification procedures	Construction permit	Land titling/ Exchange procedures	Other procedures
ATE in 2011 (<i>n</i> = 10,949)	0.041* (0.001)	0.028* (0.000)	0.037* (0.000)	0.029* (0.000)
ATE in 2012 (<i>n</i> = 11,086)	0.024* (0.002)	0.010* (0.000)	0.050* (0.000)	0.023* (0.002)
ATE in 2013 (<i>n</i> = 11,122)	-0.011 (0.311)	0.038* (0.000)	0.037* (0.000)	0.018 (0.011)
ATE in 2014 (<i>n</i> = 10,799)	-0.012 (0.299)	0.032* (0.000)	0.012* (0.000)	0.008 (0.295)
ATE in 2015 (<i>n</i> = 10,790)	0.034* (0.001)	0.011* (0.004)	0.012 (0.009)	0.008 (0.310)
Mean of control in 2011	1.73	1.76	1.54	1.83
SD of control in 2011	0.59	0.18	0.24	0.39

Table 5.
The effect of inclusion in PAPI 2010 on subdimensions of administrative procedures

Note(s): *Significant at the 0.005 level using RI, meaning that in 1,000 permutations, five or fewer had greater ATEs than the observed result. Coefficients with *p*-values (in parentheses) generated by RI clustered at the provincial level and fixed effects for the matched pairs used in the pair randomization sequences

may come from their experiences in different years, resulting in a high level of noise around each estimate due to increased variation and smaller sample sizes. On the other hand, the aggregate administrative procedure index is the non-weighted sum of the subindices,

allowing the random noise to cancel out. We thus rely on the aggregate index to demonstrate the statistical significance of the treatment effect.

6. Conclusion

By providing principals with the knowledge of agent performance, SPAs can correct the information asymmetry problem inherent in these relationships and incentivize subordinates to fulfill goals set out by their superiors. While other research has shown evidence of the effects of SPAs in democratic regimes, our study focuses on the mechanism of upward accountability in authoritarian regimes. We test our theory in Vietnam, where a quasi-meritocratic promotion system allows the information provided by PAPI – an internationally lauded governance index – to shape the behaviors of subnational leaders.

We leverage the staggered rollout of PAPI to estimate the causal effect of SPAs on governance improvement. This analysis requires a new methodology to address the dual presence of a cluster-randomized research design and sampling weights that poses countervailing effects for statistical efficiency. We resolve this problem with nonparametric RI. Ultimately, the results show that monitoring a local government’s performance significantly improves its governance and citizen satisfaction in the short run. We also find clear evidence of prioritization bias. The advantage quickly dissipates in most aspects of governance measured by PAPI as previously untreated provinces converge on the performance of those that received a head start. However, the competitive edge persists in administrative procedures, which is one particular governance area prioritized by the central government at the time.

Our conclusion concerns the gap that an informational advantage at one particular point in time creates and how persistent that gap proves to be. These findings are not meant to say that there are no overall improvements in the other five PAPI dimensions over this period. In fact, as the PAPI 2020 report shows, all areas of governance that it measures recorded sustainable improvements between 2011 and 2020, *except for public administrative procedures* (CECODES *et al.*, 2021). After PM Dung was reelected in 2011, the central government shifted policy focus away from public administration. Analysis on Google Trends shows that searches for “public administrative procedure reforms” and “Project 30” intensified markedly between 2008/2009 and 2011, before declining to their usual level of interest in later years. Incidentally, the staggered rollout of PAPI separated the provinces into two groups in 2010, which was right in the middle of this wave of public attention. Therefore, what our results demonstrate is that the informational advantage provided by an SPA coupled with the central government’s simultaneous targeted attention brought about significant improvements among the provinces endowed with the advantage as compared to those that were not. This difference remains statistically discernible for several years despite the fact that the political situation and policy discussions might have shifted.

This finding illustrates the difficulty in using performance assessments to stimulate competition and reform efforts in governance areas that are not emphasized by national leaders and have not been socialized into the practices of local leaders. Consequently, it carries critical policy implications for development interventions that seek to stimulate governance improvements through transparency initiatives. In nondemocratic systems, especially those with quasi-meritocratic promotion systems, the effectiveness of SPAs is limited by whether they cohere with central government objectives. Only central governments in these systems have the tool sets necessary to generate the material rewards and punishments that motivate reforms. Nevertheless, if the central government is reform-oriented and interested in promoting change, SPAs can have dramatic effects on the lives of citizens. Indeed, in the specific case of PAPI, the citizens welcomed the reduction in administrative procedures, which saved them valuable time and reduced their chances of being asked for petty bribes by officials.

Notes

1. Examples include the Doing Business subnational indices in China, Russia and Egypt, and The Asia Foundation's governance indices in Vietnam, Malaysia, Cambodia, Bangladesh and Myanmar. In Vietnam (the context of this paper), the Provincial Competitiveness Index (PCI) and the Provincial Administration Performance Index (PAPI) play a prominent role in informing the central government of local business environments and public administration performance, respectively.
2. Global Performance Assessments – close cousins of SPAs – such as the Doing Business Index (World Bank, 2017), Corruption Perceptions Index (Transparency International, 2017) and Freedom in the World Ranking (Freedom House, 2016) have achieved high levels of prominence in international discussions and increasing influence on policy debates in target countries.
3. For more information and the latest updates on PAPI, visit www.papi.org.vn.
4. Interview with Nguyen Thien Nhan, Chairman of the Vietnam Fatherland Front (July 15, 2016).
5. In the early years of the Vietnamese state, the VFF was an incredibly powerful organization due to the fact that nearly every citizen of the country was a member of one of its constituency organizations. Over the course of the economic reform era, however, the growth of a more complicated, multifaceted economy means that it has become possible to succeed in Vietnam without joining a mass organization. As a result, the VFF's power has diminished as evidenced by its declining share of seats in the Central Committee (Malesky, 2009) and Vietnamese National Assembly (Malesky & Schuler, 2009).
6. The Government Office is Vietnam's equivalent of the Prime Minister's Office. Phuc is now Vietnam's President.
7. Four Vietnamese search terms were entered into Google Trends: (1) "Public Administration Reform" = "Cai cach hanh chinh"; (2) "Project 30" = "De an 30"; (3) "Law on Anti-Corruption" = "Luat Phong chong tham nhung"; (4) "Grassroots Democracy Decree" = "Phap lenh thuc hien dan chu".
8. One potential approach is to take the survey sample as-is and run a regression with clustered robust standard errors. However, because this approach discards survey weights, the estimated treatment effect does not generalize to the population, which is what we are interested in. As a robustness check, we did follow this approach, running regressions with clustered robust standard errors at the province level. The direction of the sample treatment effect is the same but, as expected, its size differs from that of the population treatment effect.
9. For reference purposes, the standard deviation for Dimension 5, administrative procedures, score is 0.43 in 2015.
10. Other procedures include acquiring birth certificates, death notification, marriage certificates, ethnicity-related procedures, residency registration, housing subsidies and employment subsidies.

References

- Bell, D. (2015). *The China model: Political meritocracy and the limits of democracy*. Princeton University Press.
- Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public Goods: A political economy approach. *Journal of Public Economics*, 87(12), 2611–2637.
- Boyne, G., & Chen, A. (2007). Performance targets and public service improvement. *Journal of Public Administration Research and Theory*, 17(3), 455–477.
- Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnam Fatherland Front and United Nations Development Programme. (2010). *Towards a public administration performance index (PAPI) at the provincial level in Viet Nam*. Hanoi: Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnamese Fatherland Front and United Nations Development Programme.

- Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnam Fatherland Front and United Nations Development Programme. (2011). *The Viet Nam governance and public administration performance index (PAPI) 2010: Measuring citizens' experiences*. Hanoi: Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnamese Fatherland Front and United Nations Development Programme.
- Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnam Fatherland Front and United Nations Development Programme. (2016). *The Viet Nam governance and public administration performance index (PAPI) 2015: Measuring citizens' experiences*. Hanoi: Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnamese Fatherland Front and United Nations Development Programme.
- Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnamese Fatherland Front and United Nations Development Programme. (2021). *The 2020 Viet Nam governance and public administration performance index (PAPI 2020): Measuring citizens' experiences*. Hanoi: Centre for Community Support and Development Studies, Centre for Research and Training of the Vietnamese Fatherland Front and United Nations Development Programme.
- Chan, H. S., & Gao, J. (2012). Death versus GDP! Decoding the fatality indicators on work safety regulation in post-Deng China. *The China Quarterly*, 210, 355–377.
- Cong-Huong (2020). Chuong trinh hanh dong nang cao chi so nang luc canh tranh cap tinh (PCI) giai doan 2021–2025. *Cong thong tin dien tu tinh Bac Lieu*. Available from: <https://baclieu.gov.vn/vi/-ch%C6%B0%C6%A1ng-tr%C3%ACnh-h%C3%A0nh-%C4%91%E1%BB%99ng-n%C3%A2ng-cao-ch%E1%BB%89-s%E1%BB%91-n%C4%83ng-l%E1%BB%B1c-c%E1%BA%A1nh-tranh-c%E1%BA%A5p-t%E1%BB%89nh-pci-giai-%C4%91o%E1%BA%A1n-2021-2025>.
- Courty, P., & Marschke, G. (2004). An empirical investigation of gaming responses to explicit performance incentives. *Journal of Labor Economics*, 22(1), 23–56.
- Gandhi, J. (2004). *Political institutions under dictatorship*. New York: New York University.
- Gao, J. (2009). Governing by goals and numbers: A case study in the use of performance measurement to build state capacity in China. *Public Administration and Development*, 29(1), 21–31.
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York: WW Norton.
- Gerrish, E. (2016). The impact of performance management on performance in public organizations: A meta-analysis. *Public Administration Review*, 76(1), 48–66.
- Giang, L., Nguyen, C., & Tran, A. (2017). Does PAPI monitoring improve local governance? Evidence from a natural experiment in Vietnam. *International Journal of Development Issues*, 16(1), 85–106.
- Government of the Socialist Republic of Vietnam. (2014). *Resolution No. 19/NQ-CP dated March 18, 2014, on major tasks and solutions for improving the business environment and national competitiveness*. Hanoi: The Government of the Socialist Republic of Vietnam.
- Guo, G. (2009). China's local political budget cycles. *American Journal of Political Science*, 53(3), 621–632.
- Heinrich, C. J., & Marschke, G. (2010). Incentives and their dynamics in public sector performance management systems. *Journal of Policy Analysis and Management*, 29(1), 183–208.
- Hoai-Thu (2021). *Ha Noi quyet liet de nang cao chi so PCI trong nam 2021*. Ha Noi Moi. Available from: <https://hanoimoi.com.vn/tin-tuc/Cai-cach-hanh-chinh/1001506/ha-noi-quyet-liet-de-nang-cao-chi-so-pci-trong-nam-2021>.
- Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organizations*, 7, 24–52.
- Freedom House (2016), *Freedom in the world 2017: Populists and autocrats: The dual threat to global democracy*, Technical report, Washington, DC.

- Huu-Nghia (2021). *Dong Thap xem cac chi so PCI nhu mach mau cua nen kinh te*. Nhan Dan. Available from: <https://nhandan.vn/tin-tuc-kinh-te/dong-thap-xem-cac-chi-so-pci-nhu-mach-mau-cua-nen-kinh-te-tinh-646669/>.
- Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge: Cambridge University Press.
- Inman, R., & Rubinfeld, D. (1997). Rethinking federalism. *The Journal of Economic Perspectives*, 11(4), 43–64.
- Jackson, R. (2014). *Measuring citizen experiences: Conducting a social audit in Vietnam 2009-2013*. Princeton, NJ: Innovations for Successful Societies, Princeton University.
- Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118(3), 843–877.
- Jeong, Y. (1997). The rise of state corporatism in Vietnam. *Contemporary Southeast Asia*, 19(2), 152–171.
- Jia, R., Kudamatsu, M., & Seim, D. (2015). Political selection in China: The complementary roles of connections and performance. *Journal of the European Economic Association*, 13(4), 631–668.
- Jukic, M. (2015). *Measuring the effectiveness of government in Thailand's provinces*. DAI. Available from: <https://dai-global-developments.com/articles/measuring-the-effectiveness-of-government-in-thailands-provinces>.
- Koch, A. K., & Peyrache, E. (2011). Aligning ambition and incentives. *Journal of Law, Economics, and Organization*, 27(3), 655–688.
- Kroll, A. (2017). Can performance management foster social equity? Stakeholder power, protective institutions, and minority representation. *Public Administration*, 95(1), 22–38.
- Kung, J. K.-S., & Chen, S. (2011). The tragedy of the nomenklatura: Career incentives and political radicalism during China's great leap famine. *American Political Science Review*, 105(1), 27–45.
- Landry, P. F., Lü, X., & Duan, H. (2017). Does performance matter? Evaluating political selection along the Chinese administrative ladder. *Comparative Political Studies*, 51(8), 1074–1105.
- Lazarev, V. (2005). Economics of one-party state: Promotion incentives and support for the soviet regime. *Comparative Economic Studies*, 47(2), 346–363.
- Li, J. (2015). The paradox of performance regimes: Strategic responses to target regimes in Chinese local government. *Public Administration*, 93(4), 1152–1167.
- Magaloni, B. (2006). *Voting for autocracy: Hegemonic party survival and its demise in Mexico*. Cambridge: Cambridge University Press.
- Magaloni, B. (2008). Credible power-sharing and the longevity of authoritarian rule. *Comparative Political Studies*, 41(4-5), 715–741.
- Malesky, E. (2009). Gerrymandering-Vietnamese style: Escaping the partial reform equilibrium in a nondemocratic regime. *The Journal of Politics*, 71(1), 132–159.
- Malesky, E., & Schuler, P. (2009). Paint-by-numbers democracy: The stakes, structure, and results of the 2007 Vietnamese national assembly election. *Journal of Vietnamese Studies*, 4(1), 1–48.
- Minh-Hang (2009). *Phai don gian hoa toi thieu 30% cac quy dinh ve thu tuc hanh chinh*. baochinhpvu.vn. Available from: <https://baochinhpvu.vn/phai-don-gian-hoa-toi-thieu-30-cac-quy-dinh-ve-tthc-10219338.htm> (accessed 6 July 2022).
- Oates, W. E. (1972). *Fiscal federalism*. New York, NY: Harcourt Brace Jovanovich.
- Oxfam (2012). *Building active citizenship and accountability in Asia: Case studies from Vietnam and India*. Oxfam. Available from: <https://oxfamapps.org/fp2p/building-active-citizenship-and-accountability-in-asia-case-studies-from-vietnam-and-india/> (accessed 6 July 2022).
- Painter, M. (2005). Public administration reform in Vietnam: Foreign transplants or local hybrids? In Gillespie, J. & Nicholson, P. (Eds), *Asian socialism and legal change. The dynamics of Vietnamese and Chinese reform*, (pp. 267–287). Canberra: Australian National University.

-
- Pham, M. H. (2010). *PAR index: Piloting and next steps*. Hanoi: Ministry of Home Affairs.
- Przeworski, A., & Limongi, F. (1993). Political regimes and economic growth. *Journal of Economic Perspectives*, 7(3), 51–69.
- Reuter, O. J., & Turovsky, R. (2014). Dominant party rule and legislative leadership in authoritarian regimes. *Party Politics*, 20(5), 663–674.
- Rondinelli, D., McCullough, J., & Johnson, R. (1989). Analysing decentralization policies in developing countries: A political economy framework. *Development and Change*, 20(1), 57–87.
- Schuler, P. (2014). *Deliberative autocracy: How single-party regimes manage liberalized assemblies*. PhD thesis. San Diego: University of California.
- Schwarz, M. (2010). *Project 30: a revolution in Vietnamese governance?*. Brookings. Available from: <https://www.brookings.edu/research/project-30-a-revolution-in-vietnamese-governance/>.
- Shen, S. V. (2018). *The political pollution cycle: An inconvenient truth and how to break it*. PhD thesis. Stanford University.
- Stiglitz, J. E. (1989). Incentives, information, and organizational design. *Empirica*, 16(1), 3–29.
- Svolik, M. W. (2012). *The politics of authoritarian rule*. Cambridge: Cambridge University Press.
- Thayer, C. A. (2009). Vietnam and the challenge of political civil society. *Contemporary Southeast Asia*, 31(1), 1–27.
- The Asia Foundation. (2011). *Innovations and strengthening local economic governance in Asia*. San Francisco, CA: The Asia Foundation.
- Tiebout, C. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, 64(5), 416–424.
- Transparency International (2017). *Corruption perceptions index 2017*. Berlin: Transparency International.
- United Nations Development Programme (2001). *Modernizing governance in Viet Nam*. Hanoi: United Nations Development Programme.
- VOV. (2013). *Viet Nam va Mexico chia se kinh nghiem ve giam ngheo*. VOV. Available from: <https://vovworld.vn/vi-VN/kinh-te/viet-nam-va-mexico-chia-se-kinh-nghiem-ve-giam-ngheo-173105.vov>.
- Wallace, J. L. (2014). Juking the stats? Authoritarian information problems in China. *British Journal of Political Science*, 46(1), 11–29.
- Weingast, B. (1995). The economic role of political institutions: Market-preserving federalism and economic development. *Journal of Law, Economics, and Organization*, 11(1), 1–31.
- Whiting, S. H. (2017). Authoritarian “rule of law” and regime legitimacy. *Comparative Political Studies*, 50(14), 1907–1940.
- Wintrobe, R. (2000). *The political economy of dictatorship*. Cambridge: Cambridge University Press.
- Wong, E. (2016). Peering through the haze of China’s economic data. *The New York Times*. Available from: <https://www.nytimes.com/2016/02/27/world/asia/china-economy-outlook.html>.
- World Bank. (2017). *Doing business 2017: Equal opportunity for all*. Washington, DC: World Bank.
- Yang, D. L., Xu, H., & Tao, R. (2014). A tragedy of the nomenklatura? Career incentives, political loyalty and political radicalism during China’s great leap forward. *Journal of Contemporary China*, 23(89), 864–883.

Corresponding author

Tuan-Ngoc Phan can be contacted at: ngoc.phan@duke.edu

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com