

Statistical Capacity and Corrupt Bureaucracies

Manuel Oechslin* Elias Steiner†

— PRELIMINARY DRAFT —

October 1, 2019

Abstract

In many developing countries, economic statistics (such as the growth rate of GDP) are highly imprecise, making it difficult to evaluate economic reforms and learn “what works”. Improving economic statistics has thus become a top priority of development organizations. However, in this paper, we isolate an insidious mechanism—a type of “observer effect”—by which a push for better statistics can make matters worse. Precise statistics require the collection of data from a large number of firms. If firms suspect that detailed information, when spreading through the bureaucracy, is misused to exact bribes, they have weaker incentives to invest. As a result, the effects of reforms are muted, making it even harder to discover what works. To suppress this mechanism, efforts to improve economic statistics should be comprehensive and also include institutional measures.

JEL classification: D83, D73, F53, E01

Keywords: Partnerships in statistics, informativeness, corruption, observer effect

*University of Lucerne, Department of Economics, Frohburgstrasse 3, 6002 Lucerne, Switzerland, e-mail: manuel.oechslin@unilu.ch, phone: +41 41 229 5840. CentER, Tilburg University, The Netherlands.

†University of Lucerne, Department of Economics, Frohburgstrasse 3, 6002 Lucerne, Switzerland, e-mail: elias.steiner@unilu.ch, phone: +41 77 449 1989.

The census has nothing to do with taxation, with army or jury service [...], nor can any person be harmed in any way by furnishing the information required.
—President W.H. Taft, 1910 Census Proclamation

1 Introduction

In many developing countries, economic statistics such as the growth rate of GDP, the inflation rate, or the unemployment rate are highly unreliable. For example, in a widely noticed book, [Jerven \(2013\)](#) documented that the quality of African GDP numbers is extremely “poor”. At the same time, the World Bank’s then chief economist for Africa referred to the deficient state of African economic statistics as a “statistical tragedy” ([Devarajan, 2013](#)). It is therefore no surprise that improving developing countries’ statistics has become a priority of international organizations, among them the World Bank, the IMF, and the OECD. These organizations pursue their objective through initiatives like the “Partnership in Statistics for Development in the 21st Century” (or PARIS21), a group concerned with technical issues and the funding of data collection and processing in poor countries. More recently, the push for better statistics in developing countries has gained additional momentum through the rise of digitalization and big data.¹ Under the umbrella of the UN Global Working Group for Big Data in Official Statistics (or GWG Big Data), both developing and advanced countries exchange experiences on how to use big data to improve economic statistics.

In many ways, improvements in the precision and availability of economic statistics would be highly welcome. Reliable numbers are important for an appropriate conduct of monetary policy (e.g., [Orphanides, 2003](#)). Beyond that, considering that concepts such as “growth diagnostics” (e.g., [Rodrik, 2010](#)) and “experimentation at scale” (e.g., [Muralidharan and Niehaus, 2017](#)) have gained ground, accurate statistics are of increasing importance in the context of development policy. Growth diagnostics, for instance, is based on the notion that—when it comes to incremental reforms—which reforms work and which do not is highly context-specific, i.e., depends on a country’s economic and institutional status quo. Therefore, as [Rodrik \(2010, p. 41\)](#) puts it, growth diagnostics “emphasizes experimentation as a strategy for discovery of what works, along with monitoring and evaluation.”

A condition for meaningful monitoring and evaluation is the availability of accurate economic statistics. If the numbers are poor, evaluation may become impossible or may lead to erroneous conclusions about what works ([Manski, 2015](#)). This paper does not deny that good

¹Digitalization and big data are trends that are by no means confined to economically advanced countries. See, e.g., [Demirgüç-Kunt et al. \(2015\)](#) on the increasing importance of mobile banking in developing countries.

economic statistics have many benefits. However, focusing on development policy and growth statistics, we isolate an insidious mechanism by which in developing countries a push for better statistics can have harmful side effects. This mechanism—a type of “observer effect”—reduces the benefits or may even reverse them into net losses. Our analysis implies that efforts to improve developing countries’ statistics should not have a narrow focus on technical statistical capacity (i.e., data gathering); rather, such efforts should be comprehensive and also include institutional aspects (e.g., data confidentiality). It is key that there be a symmetry between technical statistical capacity and the strength of the institutional setting.

Our argument is based on three observations. First, strengthening technical statistical capacity to improve the accuracy of GDP statistics necessarily means collecting more data. It includes a move to a regular economic census schedule and the enlargement of the firm surveys that underlie GDP estimates between the censuses (see, e.g., [Berry et al., 2018](#); [Jerven, 2013](#), p. 26). Second, although there often are official guarantees of confidentiality of census and survey data, a large number of reports on the handling of government-collected data suggest a grave risk of confidentiality breaches² that may allow detailed data (e.g., on firm revenues) to spread widely within the bureaucracy and beyond. The reasons for this include weak governance and the prevalence of IT security holes. By one estimate from 2017, more than 80% of public sector institutions in Kenya do not even have the means to detect network intruders.³ Third, control of corruption is weaker in low-income countries (e.g., [Olken and Pande, 2012](#)), and corrupt officials use information on firm characteristics to “bribe discriminate” ([Svensson, 2003](#)), with the result that larger firms pay higher bribes (e.g., [Bai et al., 2019](#)).

Connecting these observations, a push to strengthen technical statistical capacity must arouse fears among firms of higher bribery costs: with larger surveys, each firm faces a higher chance of being sampled and, given the possibility of a confidentiality breach, a higher chance of being subject to roughly proportional bribe demands from somewhere within the bureaucracy. The expectation that, in fact, the official confidentiality (and no-harm) assurances may not hold weakens firms’ incentives to invest. But if firms invest less, the effects of economic reforms will be muted. As a result, although the improved statistics reduces the noise in growth estimates, it may become more difficult for the government to discover what works. In other words, the informativeness of policy experiments may fall rather than rise. In this case, a push to improve

²A recent example that received global attention is the illicit sale of data from India’s Aadhaar ID project, a database that covers almost the entire Indian population and includes biometric data. See, e.g., the reports in the [Guardian](#) (Jan 4, 2018), the [New York Times](#) (Apr 3, 2018), and the [Economist](#) (Dec 18, 2018)

³See [Business Daily](#) (Apr 27, 2017) on Kenya. According to a recent [Brookings Institution](#) (May 30, 2018) report on Africa, the government sector is among the top sectors impacted by confidentiality breaches.

technical statistical capacity slows down the societal learning process.

To examine the relationship between technical statistical capacity and societal learning about reforms, this paper proposes a theoretical two-period framework that features ex ante fundamental uncertainty about the effects of alternative reform options (as in [Binswanger and Oechslin, 2015, 2019](#)) and ex post measurement uncertainty in the economy’s key statistic, the output estimate. Measurement uncertainty stems from the fact that the statistical office has to base its output estimate on a random sample of firms. Within the framework, the size of this firm sample can be interpreted as a measure of the economy’s technical statistical capacity. An improvement in technical statistical capacity reduces measurement uncertainty and hence improves the accuracy of output estimates. A further key element of the framework is that it treats firms’ investment decisions as endogenous and allows for the possibility of firms being subject to bribe demands by bureaucrats. Specifically, we assume that firms sampled by the statistical office face a positive probability that they will be confronted with (additional) bribe demands that amount to a fixed proportion of their current revenue.⁴

In this framework, what are the consequences of an exogenous improvement in technical statistical capacity? Holding constant firms’ investments, a fall in measurement uncertainty permits a more reliable assessment of whether an implemented reform boosts output or whether the government should pursue adjustments to make the reform work; as a result, the learning process speeds up. However, firms’ investments do not stay constant when technical statistical capacity improves: a larger sample implies that each individual firm faces a higher probability of being sampled and hence a higher risk of being confronted with bribe demands; as a result, firms scale back their investments, a response that has direct negative consequences for economic performance. But even more importantly, with smaller investments, economic reforms have a smaller effect on output—which, in turn, makes it more difficult for the government to determine whether an implemented reform works or needs adjustment.

So an improvement in technical statistical capacity has two opposing effects on the speed of the societal learning process and hence economic performance. A key implication of our framework is that—if control of corruption is sufficiently weak—there is a hump-shaped relationship between technical statistical capacity and economic performance: increasing the firm sample helps initially but reduces expected output beyond some critical threshold. In other words, although sampling is costless and the government interested in learning about reforms, the optimal sample size is strictly smaller than the total number of firms. Corruption, by

⁴While the baseline setup is kept parsimonious, we present two obvious extensions in the appendix. One of them allows for misreporting by firms, a potential issue when bribe demands are influenced by reported revenue.

impairing the government’s ability to identify the consequences of its reform decisions, retards economic growth by slowing down the learning process about what works.⁵

Over the past few years, a growing literature on the quality of economic statistics in developing countries has emerged. In many poor countries, the quality is low, implying a large potential for improvements in the timeliness and precision of economic indicators (e.g., [Devarajan, 2013](#); [Jerven, 2013](#); [Kiregyera, 2015](#); [Sandefur and Glassman, 2015](#)).⁶ A large number of contributions emphasize the link between the quality of economic statistics and policy making. [Rodrik \(2010\)](#), for instance, stresses the importance of high quality data for evidence-based development policy. [Manski \(2015\)](#) worries that imprecise estimates may lead to bad policy decisions if policy-makers fail to account for measurement error. [Binswanger and Oechslin \(2015\)](#) argue that better statistics—by making evaluations of past policy changes more reliable—could reduce disagreement and promote economic reforms. More in line with the present paper, [Binswanger and Oechslin \(2019\)](#) identify adverse consequences of better statistics in electoral democracies. Although the present paper also uses a model of active policy learning, it differs significantly from the former. Here, we relate learning about policies to corruption and explicitly model firms’ investment choices in this context. Thereby, we connect with the literature on corruption and firms, amongst others with firm-level studies documenting that corruption constrains the growth aspirations and the advancement of firms (e.g., [Fisman and Svensson, 2007](#); [Estrin et al., 2013](#); [Freund et al., 2016](#); [Colonnelli and Prem, 2017](#)).

The rest of this paper is organized as follows. The upcoming section presents motivating evidence. In Section 3, we describe the theoretical setup. Section 4 solves for the equilibrium, assuming a given level of technical statistical capacity. Section 5 derives the optimal capacity level and discusses the harmful role of corruption. Section 6 concludes.

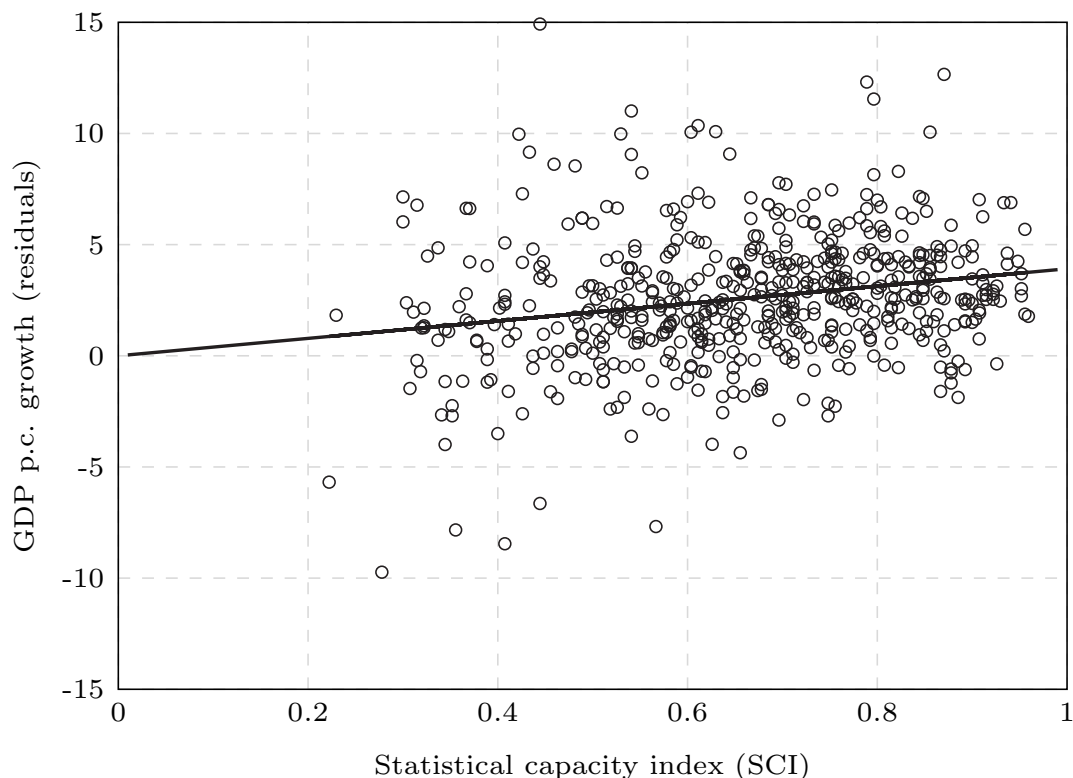
2 Motivating Evidence

In this section, we present motivating evidence on the relationship between the quality of economic statistics and economic performance. Thereby, we also consider the moderating role of corruption. To capture the quality of economic statistics, we use the World Bank’s Statistical Capacity Index (SCI). The SCI is available for 153 developing and emerging economies at yearly

⁵There is a rich literature on economic policy channels through which corruption may affect growth. An influential part of this literature focuses on channels emphasizing the impact of corruption on taxation and public goods (see, e.g., [Aghion et al., 2016](#)). The policy learning channel has not been investigated so far.

⁶There is a related literature in macroeconomics, real-time data analysis, that documents properties of data revisions and explores how such revision matter, e.g., for monetary policy analysis. See [Croushore \(2011\)](#).

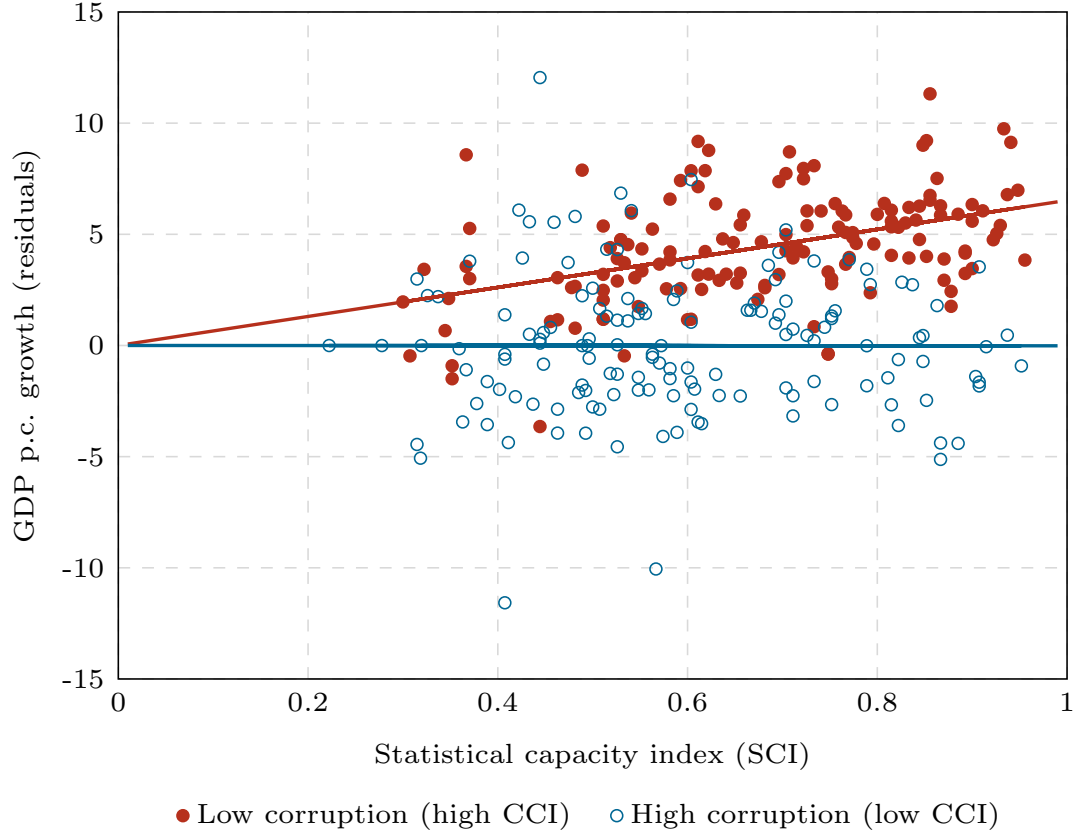
Figure 1: Statistical capacity and growth, full sample, 2005-2016



Note: This figure shows a partial residual plot. The underlying linear OLS regression relates the average annual growth rate of real GDP p.c. to a constant, the SCI, the log of real GDP p.c., and period fixed effects. The value of the coefficient on the SCI (which is also the slope of the fitted line) is 3.913 (p -value: 0.000).

frequency. We have recoded the index so that it ranges from 0 to 1, where 1 indicates maximum statistical capacity (the original range is 0 to 100). The index measures the extent to which a country's statistical system adheres to international technical standards deemed essential for the quality of economic data. We use the growth rate of real GDP p.c. (PPP, constant 2011 I\$) to capture economic performance and the World Bank's Control of Corruption Index (CCI) as a measure for (the absence of) corruption. The CCI is a corruption perception index that is concerned with the exercise of public power for private gain. Again, we have recoded the index so that it ranges from 0 to 1, where 1 indicates maximum control of corruption. Our dataset includes 146 countries and covers the period from 2005 to 2016. In the figures below, we use observations averaged over periods of three years (2005-07, 2008-10, 2011-2013, and 2014-16).

Figure 2: Statistical capacity and growth, two subsamples, 2005-2016



Note: This figure shows a partial residual plot for countries with an average CCI score belonging to the top (“low corruption”) or bottom (“high corruption”) 25% of the distribution (145 observations each). The underlying linear OLS regressions relate the average annual growth rate of real GDP p.c. to a constant, the SCI, the log of real GDP p.c., and period fixed effects. The values of the coefficients on the SCI (which are also the slopes of the fitted lines) are -0.015 (p -value: 0.994) for the high-corruption subsample and 6.527 (p -value: 0.000) for the low-corruption subsample. The difference in the two coefficients is statistically significant at the 5% level.

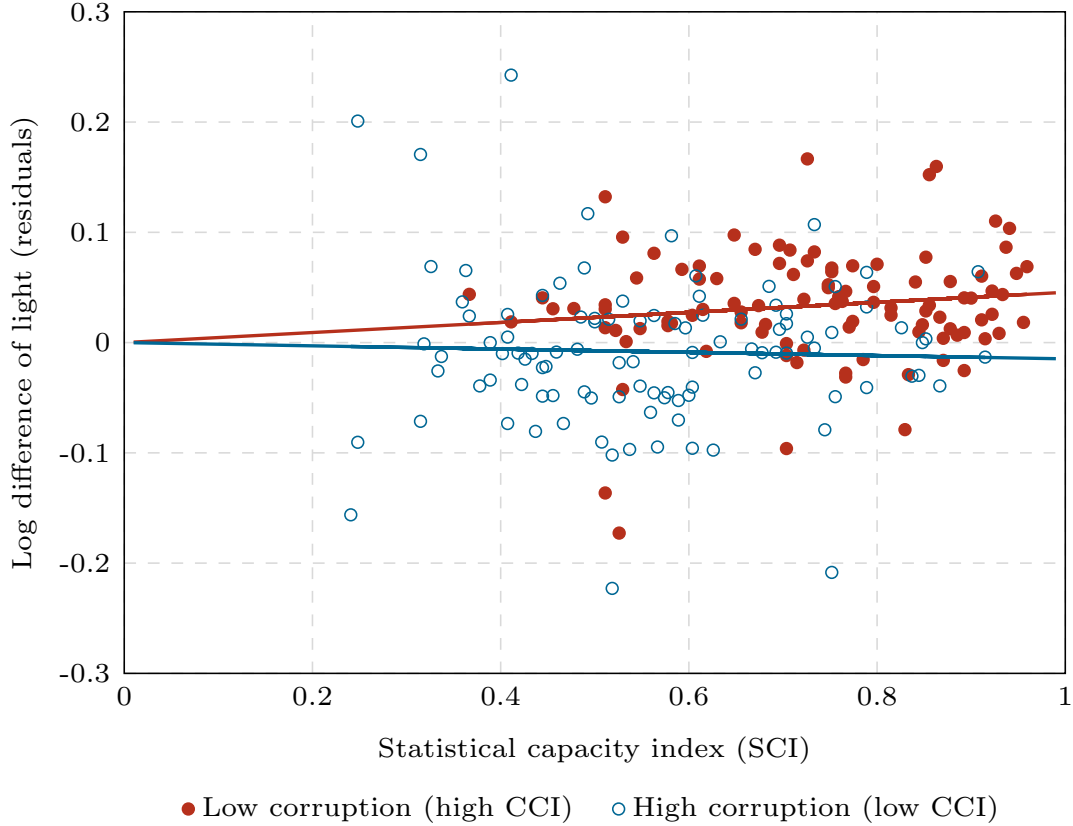
The full sample includes 556 observations.⁷

Figure 1 shows a partial residual plot that illustrates the correlation between the residual growth rate of real GDP p.c. and statistical capacity.⁸ We see a significant positive relationship: an increase in the SCI of one standard deviation (0.16) is associated with a rise in real GDP p.c.

⁷The data were retrieved on July 4, 2018, from <https://data.worldbank.org/>. We dropped as outliers all observations that belong to either the bottom or top 0.5% in the distribution of GDP p.c. growth rates.

⁸Besides SCI, the underlying OLS regression includes as explanatory variables the log of real GDP p.c. (to control for convergence) and period fixed effects (to control for period-specific shocks common to all countries).

Figure 3: Statistical capacity and nighttime light intensity, two subsamples, 2005-2013



Note: This figure shows a partial residual plot for countries with an average CCI score belonging to the top (“low corruption”) or bottom (“high corruption”) 25% of the distribution (101 and 102 observations, respectively). The underlying linear OLS regressions relate the log difference of mean light intensity to a constant, the SCI, the log of mean light intensity, and period fixed effects. The values of the coefficients on the SCI (which are also the slopes of the fitted lines) are -0.015 (p -value: 0.771) for the high-corruption subsample and 0.046 (p -value: 0.215) for the low-corruption subsample.

growth of 0.6 percentage points. Given an interquartile range of real GDP p.c. growth of just 3.4 percentage points, this is a sizable correlation. Figure 1 is based on the full sample and does not account for cross-country differences in corruption. The role of corruption is highlighted in Figure 2. Again, the figure shows a partial residual plot, but it separately considers two disjunct subsamples of the full sample. One of them contains all countries with an average CCI score belonging to the top 25% of the distribution (“low corruption”) and the other one all countries with an average CCI score belonging to the bottom quartile (“high corruption”). Figure 2

suggests that the positive correlation that is apparent in Figure 1 is driven by observations from low-corruption countries. While there is a significant positive relationship between the growth rate of real GDP p.c. and statistical capacity in the low-corruption subsample, no such relationship emerges among high-corruption countries.

The pattern documented in Figure 2 is fairly robust to a number of modifications. It remains mostly unaffected when we use larger subsamples of low- and high-corruption countries.⁹ It is also robust to the inclusion of country fixed effects in the underlying regressions (though the p -value on the estimate of the coefficient on SCI in the low-corruption subsample rises from virtually 0 to 0.15). Finally, given the concerns regarding GDP data quality emphasized in the introduction, we were also using the (log) change in nighttime light intensity as a proxy for economic performance (see Henderson et al., 2012). The source of the light data is Hodler and Raschky (2014), who aggregated the georeferenced raw data to ADM2 administrative levels (from where we aggregated it to the country level).¹⁰ The data are scaled from 0 to 63, where a larger number reflects more intense nighttime lights. The data are available up to the year 2013 only, which leaves us with 399 observations from 134 countries. Overall, we find that the swap of GDP for light data does not change the basic pattern documented in Figure 2: there tends to be a positive relationship between economic performance and statistical capacity among low-corruption countries, while no such relationship appears among high-corruption countries. However, the differences between low- and high-corruption countries are smaller and more sensitive to the threshold applied. In Figure 3, the threshold is 25%.

It is clear that one cannot draw any firm conclusions from the above results. However, they do draw attention to the fact that the correlation between economic performance and statistical capacity—while being positive overall—is not uniform but dependent on corruption. The framework developed below offers an explanation for this pattern.

⁹The 25%-threshold (top/bottom) is chosen for visual clarity. The general conclusion stated at the end of the preceding paragraph remains valid when we gradually move to larger thresholds, up to the median.

¹⁰The raw data come from the National Oceanic and Atmospheric Administration (NOAA) (2014). To aggregate it, Hodler and Raschky (2014) use the GADM database of Global Administrative Areas (2012).

3 The Model

3.1 Firm Output and Economic Policy

We consider a two-period economy that is populated by $N > 0$ firms that produce a uniform final good. Total output by firm $i \in \{1, \dots, N\}$ in period $t \in \{1, 2\}$ is given by

$$z_{it} = y_{it} + \zeta_{it}, \tag{1}$$

where y_{it} and ζ_{it} refer to output produced using a “modern” and a “traditional” technology, respectively. The simultaneous use of two technologies by the same firm may be the result of firms producing in multiple locations with different degrees of suitability for the technologies. The modern technology is represented by the production function

$$y_{it} = y(A_t, x_{it}) = A_t x_{it}^\alpha, \tag{2}$$

where $0 < \alpha < 1$. A_t captures the level of productivity of the modern technology and x_{it} refers to a firm-specific investment whose per-unit cost is normalized to 1. Productivity is affected by economic policy, P_t . Following [Binswanger and Oechslin \(2019\)](#), we assume that $P_t \in \{-1, 0, 1\}$, where -1 and 1 denote two alternative “reform policies” and 0 reflects the “status-quo policy”. The impact of policy on productivity is described by

$$A_t = 1 + \sqrt{\gamma} P_t S, \tag{3}$$

where $0 < \gamma < 1$ reflects the economic significance of the reform and $S \in \{-1, 1\}$ captures the unobserved and invariable “state of the world” that materializes prior to the start of the economy. Equations (2) and (3) together imply that a reform policy is beneficial (harmful) if its sign is the same as (is different from) the sign of the state. S takes each of its two possible values with probability $1/2$. This specific value is chosen for analytical convenience and not important for our argument. What matters is that there is some uncertainty as to whether a particular reform alternative has a positive or a negative effect on output.

The sole purpose of having a traditional technology is to generate idiosyncratic variation in firm level output that is independent of policy.¹¹ So the modeling is parsimonious: ζ_{it} is a continuous i.i.d. random variable with support on $[0, \infty)$, mean $\chi > 0$ and variance σ . For a reason that will become clear below, we assume that the distribution of ζ_{it} has a light left tail.

¹¹The idea is that the modern technology involves a high degree of specialization and hence crucially depends on government policies (such as those related to contract enforcement), while the traditional technology is very basic so that it can be operated without government services. See, e.g., [Acemoglu et al. \(2007\)](#).

Moreover, to ensure the model's scale invariance regarding N , we impose

$$\sigma = \theta N, \tag{4}$$

where $\theta > 0$. As will become clear below, the parameter θ governs the volatility of average traditional technology output. Finally, in what follows, we use an asterisk (*) to mark values reflecting optimal firm choices. For instance, we use y_{it}^* and $z_{it}^* = y_{it}^* + \zeta_{it}$ to denote firm i 's optimal modern technology output and optimal total output, respectively.

3.2 Informational Constraints and Output Estimation

Each firm i observes z_{it}^* but does not directly observe y_{it}^* or ζ_{it} within the same period, reflecting frictions in the gathering and/or processing of detailed information within firms. Because of this, the implementation of a reform policy in period 1 does not put firms in a position that would allow them to unambiguously infer state S within that very period. However, over time, firms become able to separately observe y_{i1}^* and ζ_{i1} and hence—given $P_1 \neq 0$ —identify the state. In particular, firm i learns (y_{i1}^*, ζ_{i1}) just before it chooses its level of investment in period 2 (an assumption that simplifies our analysis without affecting its substance).

Prior to the possible identification of state S in period 2, all agents can turn to information provided by the statistical office to update their belief about S : the statistical office is endowed with the capacity to collect firm-level data on z_{it}^* , which it uses to calculate and publish estimates of average total firm output at the end of each period. The estimates are based on a random sample of $n \leq N$ firms, where n is determined before the start of the economy. The ratio $p \equiv n/N$ is an obvious measure of *technical statistical capacity* in our simple economy. The activities of the statistical office (data collection, processing, and publication) are costless. The statistical office's estimate of average total firm output for period t is given by

$$Z_t^p \equiv \frac{1}{pN} \sum_{i=1}^{pN} z_{it}^* = \left(\frac{1}{pN} \sum_{i=1}^{pN} y_{it}^* \right) + \zeta_t^p, \tag{5}$$

where

$$\zeta_t^p \equiv \frac{1}{pN} \sum_{i=1}^{pN} \zeta_{it}. \tag{6}$$

Assuming that pN is sufficiently large, the Lindeberg-Lévy CLT implies that the distribution of $\sqrt{pN}(\zeta_t^p - \chi)$ is closely approximated by $N(0, pN)$. Accordingly, we impose

$$\zeta_t^p \sim N(\chi, \theta/p). \tag{7}$$

If $P_1 \neq 0$, the publication of Z_t^p at the end of period 1 allows all agents to compute the ex-post probability that the implemented reform alternative is the beneficial one:

$$r(Z_1^p) \equiv \Pr [P_1 = S | Z_1^p]. \quad (8)$$

To compute $r(Z_1^p)$, all agents rely on Bayes' rule.¹²

3.3 The Bureaucracy and the Government

Firm-level data collected by the statistical office may be misused: in both periods, there is a probability $\pi > 0$ of a confidentiality breach that puts the data into the hands of a “corrupt” government official who can extract bribes from those firms he has got information on. Following [Svensson \(2003\)](#) and [Fisman and Svensson \(2007\)](#), we may think of an official outside the statistical office whose power to collect bribes derives from his discretion in the application and enforcement of complex regulations. To get hold of the confidential firm data, the official may work with accessories inside the statistical office or may find a way to intrude the office's IT system. Yet the official's power is limited: firms can fend off bribe demands by incurring a cost amounting to a fraction $0 < \hat{\beta} < 1$ of total firm output (but prefer paying the bribe to fending the demands off if the costs are the same).¹³ So, provided that a confidentiality breach occurs in period t , any given firm i sampled in that period will face bribe demands of size $\hat{\beta}z_{it}$. We will refer to $\beta \equiv \pi\hat{\beta}$ as the bureaucracy's *vulnerability to corruption*.

The government determines statistical capacity and is in charge of economic policy. We consider a reform-minded government that has inherited a bureaucracy whose vulnerability to corruption cannot be changed within the time horizon of the model and hence must be taken as exogenous. The government's objective is to maximize the expected lifetime total output by the representative firm i . Formally, its objective function is given by

$$V = E_1 \{z_{i1}^* + z_{i2}^*\}, \quad (9)$$

where the expectation in equation (9) is formed at the beginning of period 1. So our analysis of optimal technical statistical capacity is carried out in a “favorable” setting in which the government does not pursue any special interests but aims at maximizing aggregate output and in which the operations of the statistical office are free of charge.

¹²There is a chance that the realization of ζ_{i1} is so close to 0 so that $y_{i1}^*|_{P_1 \neq S} + \zeta_{i1} < y_{i1}^*|_{P_1 = S}$, in which case firm i —and the statistical office if i is sampled—can infer S in period 1 if $P_1 \neq S$. But since we assume that the distribution of ζ_{i1} has a light left tail, this chance is so small that it can be ignored in our analysis.

¹³[Bai et al. \(2019\)](#) argue that competition among sub-national jurisdictions to attract firms keeps bribe extraction in check: if the bribe a corrupt official asks for is “too high”, the affected firm will move away.

3.4 Time Line

The timing of actions is as follows. Prior to the start of the economy, Nature determines the unobserved state of the world S and the government chooses statistical capacity p .

In the first period, the government sets P_1 ; observing the government’s policy decision, all firms choose x_{i1} ; the statistical office draws the random firm sample, thereby observing the government’s choice of p ; Nature determines $\{\zeta_{i1}\}_{i=1}^N$; the statistical office collects the data and—if there is a confidentiality breach—the firms belonging to the sample are asked for bribes; the statistical office publishes Z_1^p and—if $P_1 \neq 0$ —all actors compute $r(Z_1^p)$.

In the second period, the government sets P_2 ; all firms learn (y_{i1}^*, ζ_{i1}) and—if $P_1 \neq 0$ —infer state S ; taking P_2 and the available information on the state into account, all firms choose x_{i2} . From this point onwards, the sequence of actions is identical to that in period 1.

3.5 Extensions

In the baseline setup proposed above, it is simply assumed that firms that are not part of the statistical office’s sample in period t will not be asked for bribes in that period. Appendix B presents an extension of the baseline setup in which such differential treatment of sampled and non-sampled firms can be the result of optimizing behavior on the part of the official. The logic is straightforward. When directed at sampled firms, the official is able to “bribe discriminate” (Svensson, 2003), i.e., to adjust his demands such that they match each individual firm’s maximum willingness to pay. However, when approaching non-sampled firms, the official has to ask for a uniform bribe since firm output is private information. As a result, the expected payoff from approaching a random non-sampled firm is small relative to that of approaching a sampled firm. So, if approaching firms comes at an appropriate effort cost, it is optimal to approach sampled firms but to leave non-sampled firms alone.

Another important assumption in the above baseline setup is that the statistical office is in a position to accurately observe the output of those firms that are part of its random sample. Appendix C presents a different extension of the baseline setup in which the statistical office is in a weaker position: instead of observing output levels, the office has to work with output data as reported by the firms. This poses a potential problem: sampled firms, anticipating that they may be subject to proportional bribe demands, have an incentive to misreport. However, any misreporting carries a fine—provided it is detected. We show that the key implications of the baseline setup are robust to this modification. The logic is again straightforward. The statistical office, understanding the firms’ incentives to misreport, can infer the actual

output levels from the (mis-)reported ones. Put differently, when calculating its average output estimate, the office applies an appropriate “correction” to the (mis-)reported data. As a result, the average output estimate is as accurate as in the baseline setup.

4 Equilibrium Economic Policy

4.1 Input Choice

Before going backwards through the sequence of policy choices, we consider firms’ investment decisions. In period $t \in \{1, 2\}$, firm i solves the maximization problem

$$\begin{aligned} \max_{\{x_{it}\}} \left\{ p \left[\pi(1 - \hat{\beta}) E_t^i \{y(A_t, x_{it}) + \zeta_{it}\} + (1 - \pi) E_t^i \{y(A_t, x_{it}) + \zeta_{it}\} \right] \right. \\ \left. + (1 - p) E_t^i \{y(A_t, x_{it}) + \zeta_{it}\} - x_{it} \right\}, \end{aligned} \quad (10)$$

where $E_t^i\{\cdot\}$ refers to the expectation formed by the firm just before it chooses x_{it} . The objective function in problem (10) reflects that, with probability p , firm i is sampled by the statistical office, in which case it faces a chance of π of being confronted with bribe demands that amount to a fraction $\hat{\beta}$ of its output. Problem (10) can be algebraically simplified to

$$\max_{\{x_{it}\}} \left\{ (1 - p\beta) (E_t^i \{A_t\} x_{it}^\alpha + \chi) - x_{it} \right\}. \quad (11)$$

Since $0 < \alpha < 1$, the objective function in maximization problem (11) is a strictly concave function of $x_{it} \in [0, \infty)$. The function’s maximizer is given by

$$x_{it}^* = [\alpha(1 - p\beta) E_t^i \{A_t\}]^{1/(1-\alpha)}, \quad (12)$$

Using production function (2), we can calculate firm i ’s optimal modern technology output as

$$y_{it}^* = A_t [\alpha(1 - p\beta) E_t^i \{A_t\}]^{\alpha/(1-\alpha)}. \quad (13)$$

4.2 Second Period

The final decisions of interest to be taken in period 2 are those by the firms on second-period investment. When making their decisions, the firms have just learned about (y_{i1}^*, ζ_{i1}) . As a result, if $P_1 \neq 0$, they can infer $S \in \{-1, 1\}$ from (y_{i1}^*, ζ_{i1}) ; moreover, having observed P_2 , they can identify A_2 with certainty (equation 3). Otherwise, if $P_1 = 0$, firms still believe that state S takes each of its possible values with probability $1/2$; as a result, we can conclude that $E_2^i \{A_2\} = 1$ irrespective of the choice of P_2 . To summarize:

$$E_2^i \{A_2\} = \begin{cases} 1 & : P_1 = 0 \\ A_2 & : P_1 \neq 0 \end{cases}. \quad (14)$$

The first decision to be taken in period 2 is that by the government on second-period policy, $P_2 \in \{-1, 0, 1\}$. Objective function (9) implies that, at this point in time, the government wants to maximize $E_2 \{z_{i2}^*\}$, where the expectation is formed at the beginning of period 2. First assume that the status-quo policy has been implemented in period 1 ($P_1 = 0$). For this case, equation (14) implies $E_2^i \{A_2\} = 1$. As a result, we obtain

$$E_2 \{z_{i2}^*(P_2)\} = E_2 \{A_2(P_2)\} [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} + \chi. \quad (15)$$

Again, since the beliefs about state S cannot be updated in this case, we have $E_2 \{A_2(P_2)\} = 1$ irrespective of the particular choice of P_2 . As a result, the government is indifferent between the three policy options. Without loss of generality, we henceforth assume that the government decides to keep the status-quo policy in place:

$$P_2|_{P_1=0} = 0. \quad (16)$$

Now suppose that a reform policy has been implemented in period 1 ($P_1 \neq 0$). In this case, taking into account equations (13) and (14), we obtain

$$E_2 \{z_{i2}^*(P_2)\} = E_2 \left\{ [A_2(P_2)]^{1/(1-\alpha)} \right\} [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} + \chi. \quad (17)$$

The expectation in equation (17) is now based on $r(Z_1^p)$, the ex-post probability that the first-period reform is beneficial. Therefore:

PROPOSITION 1 *Suppose $P_1 \neq 0$. Then, in order to maximize $E_2 \{z_{i2}^*(P_2)\}$, the government chooses P_2 according to*

$$P_2|_{P_1 \neq 0} = \begin{cases} P_1 & : r(Z_1^p) \geq 1/2 \\ -P_1 & : r(Z_1^p) < 1/2 \end{cases}. \quad (18)$$

Proof. See Appendix A. ■

4.3 First Period

If a reform policy has been implemented, the final activity in period 1 is the formation of the posterior belief, $r(Z_1^p) \equiv \Pr [P_1 = S | Z_1^p]$. Since, from an ex-ante perspective, state S takes each of its two possible values with probability 1/2, it follows that $E_1^i \{A_1\} = 1$ irrespective of the value of P_1 . As a result, equation (13) implies that the optimal modern-technology output chosen by firm i in period 1 is given by

$$y_{i1}^* = A_1 [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)}. \quad (19)$$

Given this, and considering equations (5) and (7), Z_1^p follows a normal distribution, the mean of which being a function of state S and policy P_1 :

$$Z_1^p \sim N \left\{ A_1 [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} + \chi, \theta/p \right\}, \quad (20)$$

where $A_1 = 1 + \sqrt{\gamma}$ if $P_1 = S$ and $A_1 = 1 - \sqrt{\gamma}$ if $P_1 = -S$. Since each of these two possibilities materializes with probability 1/2, Bayes' rule implies

$$r(Z_1^p) \equiv \frac{(1/2) \cdot f(Z_1^p | P_1 = S)}{(1/2) \cdot f(Z_1^p | P_1 = S) + (1/2) \cdot f(Z_1^p | P_1 = -S)}, \quad (21)$$

where $f(Z_1^p | \cdot)$ denotes the corresponding normal density. Using functional forms, we obtain

$$r(Z_1^p) = \left\{ 1 + \exp \left[2 \frac{\gamma}{\sigma/n} \left(\hat{Z}_1^p - \chi \right) \left(\hat{Z}_1^p - Z_1^p \right) \right] \right\}^{-1}, \quad (22)$$

where

$$\hat{Z}_1^p = [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} + \chi. \quad (23)$$

According to equation (22), $r(Z_1^p)$ is a strictly increasing function of Z_1^p , rising from 0 (if $Z_1^p \rightarrow -\infty$) to 1/2 (if $Z_1^p = \hat{Z}_1^p$) to 1 (if $Z_1^p \rightarrow \infty$). In combination with Proposition 1, this implies that $P_2 = P_1$ if $Z_1^p \geq \hat{Z}_1^p$ and $P_2 = -P_1$ otherwise.

Moving backwards to the firms' decision, we note that x_{i1}^* and y_{i1}^* are given by equations (12) and (13), respectively, with $E_1^i \{A_1\} = 1$ irrespective of the value of P_1 .

The first decision to be taken in period 1 is that by the government on first-period policy. To inform this decision, the government compares the value of its objective function, $V = E_1 \{z_{i1}^* + z_{i2}^*\}$, under the status quo to the value under any of the two reform alternatives. According to equation (16), $P_1 = 0$ implies $P_2 = 0$. As a result, if the government opts for the status quo in period 1 ($P_1 = 0$), we obtain $A_1 = A_2 = 1$. From this, it follows that the expected lifetime total output by the representative firm i is given by

$$E_1 \{z_{i1}^* + z_{i2}^* | P_1 = 0\} = 2 [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} + 2\chi. \quad (24)$$

Because of the symmetric setup, the government is indifferent between the reform alternatives -1 and 1 . Without loss of generality, we henceforth assume that $P_1 = 1$ if the government decides to abandon the status quo. In this case, P_2 is determined by equation (18) and the expected lifetime total output by the representative firm i is given by

$$E_1 \{z_{i1}^* + z_{i2}^* | P_1 = 1\} = \left[1 + E_1 \left\{ [A_2(P_2)]^{1/(1-\alpha)} \right\} \right] [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} + 2\chi. \quad (25)$$

Moreover:

LEMMA 1 *Suppose the government opts for a reform policy in period 1 (e.g., $P_1 = 1$). Then,*

$$E_1 \left\{ [A_2(P_2)]^{1/(1-\alpha)} \right\} = \hat{A}^l + \Pr[P_2 = S] (\hat{A}^h - \hat{A}^l) > 1, \quad (26)$$

where $\Pr[P_2 = S]$ denotes the probability that the government will choose the beneficial reform policy in period 2 and

$$\hat{A}^l \equiv (1 - \sqrt{\gamma})^{1/(1-\alpha)} \quad \text{and} \quad \hat{A}^h \equiv (1 + \sqrt{\gamma})^{1/(1-\alpha)}. \quad (27)$$

Proof. See Appendix A. ■

The results established so far lead to the following conclusion:

PROPOSITION 2 *In period 1, the government prefers reform (e.g., $P_1 = 1$) to the status quo. In period 2, the government's policy choice is described by equation (18).*

Proof. The first statement of the proposition follows from equations (24) and (25) and Lemma 1. The second statement follows from the first and Proposition 1. ■

In period 1, there are two factors that make the government prefer a reform policy to the status-quo. First, if a reform policy is implemented, the government gains information about “what works”; this, in turn, allows for a better-informed policy decision in period 2. Second, output produced using the modern technology is convex in A_2 ; as a result, taking a “symmetric risk” is preferred to obtaining the expected value with certainty.

5 Statistical Capacity

5.1 Optimal Statistical Capacity

Besides economic policy, the government determines technical statistical capacity, p , with a view to maximizing the expected lifetime total output by the representative firm i . According to equations (25) and (26), an important magnitude in the government's optimization problem is the probability with which it will implement the beneficial reform policy in period 2.

PROPOSITION 3 *At the beginning of period 1, i.e., at the moment when the government decides to implement a reform policy (Proposition 2), the probability that the implemented second-period reform is in fact the beneficial one, $\Pr[P_2 = S]$, is given by*

$$I(C; \gamma, \theta) \equiv \Phi \left(\sqrt{\gamma/\theta} \cdot C(p; \alpha, \beta) \right) \quad (28)$$

where $\Phi(\cdot)$ denotes the distribution function of the standard normal distribution and

$$C(p; \alpha, \beta) \equiv \sqrt{p} [\alpha (1 - p\beta)]^{\alpha/(1-\alpha)} > 0. \quad (29)$$

Proof. See Appendix A. ■

In what follows, we will call I the “informativeness of policy experimentation”.¹⁴ As can be seen from equation (28), informativeness depends on the one hand on variables unrelated to the statistical office: other things equal, if a reform is more significant (higher γ), or if aggregate traditional technology output is less volatile (lower θ), informativeness is higher. On the other hand, informativeness is influenced by the statistical office. In equation (28), $C(p; \alpha, \beta)$ captures the entirety of channels by which the statistical office affects informativeness. For this reason, we will call $C(p; \alpha, \beta)$ a measure of “comprehensive statistical capacity”. It is immediately apparent that the effect of technical statistical capacity on comprehensive statistical capacity is ambiguous if the bureaucracy’s vulnerability to corruption is not identical to zero ($\beta > 0$). This reflects that firms—observing a positive relationship between p and expected bribe demands—reduce investment in response to a rise in technical statistical capacity (equation 12). The parameter α enters $C(p; \alpha, \beta)$ because it governs the elasticity of investment with respect to expected bribe demands. $C(p; \alpha, \beta)$ has the following important properties:

LEMMA 2 *Comprehensive statistical capacity $C(p; \alpha, \beta)$ is a function of p on $[0, 1]$ that has a unique maximizer, $p^* \in (0, 1]$. Moreover, $C(p; \alpha, \beta)$ is strictly concave on $[0, p^*)$.*

Proof. See Appendix A. ■

What level of technical statistical capacity maximizes comprehensive statistical capacity? The answer depends to a large degree on the bureaucracy’s vulnerability to corruption:

PROPOSITION 4 *If the bureaucracy is sufficiently vulnerable to corruption, the level of technical statistical capacity that maximizes comprehensive statistical capacity $C(p; \alpha, \beta)$ —and hence informativeness $I(C; \gamma, \theta)$ —is strictly less than 1. In formal terms: if*

$$\beta > \frac{1 - \alpha}{1 + \alpha}, \tag{R1}$$

we obtain

$$p^* = \frac{1 - \alpha}{1 + \alpha} \frac{1}{\beta} < 1, \tag{30}$$

where we use p^* to denote the maximizer of $C(p; \alpha, \beta)$.

Proof. See Appendix A. ■

¹⁴The complementary probability of $P_2 = S$ is equal to the sum of the probabilities of a type-I error ($P_1 = S$ is rejected although true) and of a type-II error ($P_1 = S$ is not rejected although false).

A rise in technical statistical capacity has opposing effects on comprehensive statistical capacity and hence on the extent to which the estimate of average firm output, Z_1^p , is informative for the choice of policy in period 2. On the positive side, a rise in technical statistical capacity reduces the variance of the exogenous component of Z_1^p (equation 7); other things equal, this helps informativeness. On the negative side, for any individual firm, a rise in technical statistical capacity implies a higher chance of being confronted with bribe demands. As a result, firms respond by reducing investment (equation 12)—which dampens the impact of reforms on firm output. Other things equal, this harms informativeness. The strength of the negative effect rises in the bureaucracy’s vulnerability to corruption. If β is sufficiently large, the negative effect starts to dominate at a strictly interior level of statistical capacity.

However, comprehensive statistical capacity—and the associated probability of choosing the beneficial reform policy in period 2—is not the only variable the government has to consider when determining the level of p that maximizes V , the expected lifetime total output by the representative firm. The negative static effect of technical statistical capacity on firm investment (and hence firm output) must be accounted for, too. Therefore:

PROPOSITION 5 *Suppose that the bureaucracy is sufficiently vulnerable to corruption so that condition (R1) holds. Then, the equilibrium level of technical statistical capacity (i.e., the level that maximizes lifetime total output of the representative firm, V) is strictly less than the level that maximizes comprehensive statistical capacity $C(p; \alpha, \beta)$. In formal terms:*

$$p^{**} < p^* = \frac{1 - \alpha}{1 + \alpha} \frac{1}{\beta} < 1, \quad (31)$$

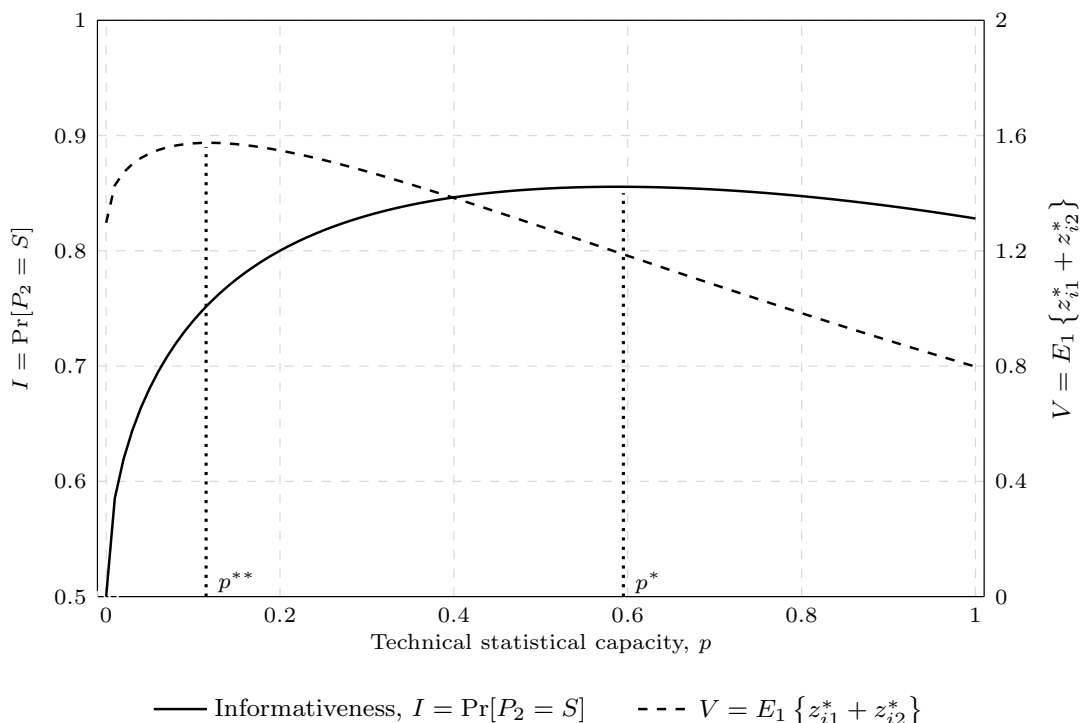
where we use p^{**} to denote equilibrium technical statistical capacity.

Proof. See Appendix A. ■

Figure 4 illustrates how $I(C; \gamma, \theta)$ and $V = E_1 \{z_{i1}^* + z_{i2}^*\}$ depend on technical statistical capacity, assuming that condition (R1) holds. Both curves in the figure are hump-shaped. Proposition 5 predicts that V peaks at a strictly lower level of technical statistical capacity than C does. As a result, the peak of I must lie to the right of the peak of V .

In Figure 4, p^{**} takes a particularly low value. So equilibrium technical statistical capacity is weak, as is technical statistical capacity in many developing countries. The figure thus conveys the message that real-world instances of low technical statistical capacity should be interpreted with care. In addition to, or instead of, reflecting neglect or a lack of resources and expertise, keeping technical statistical capacity at a low level may be the best response by a government that is confronted with an intractably adverse institutional setting.

Figure 4: Technical statistical capacity, informativeness, and economic performance



For the simulation we have chosen the following parameter values: $\alpha = 0.7$, $\beta = 0.3$, $\gamma = 0.5$.

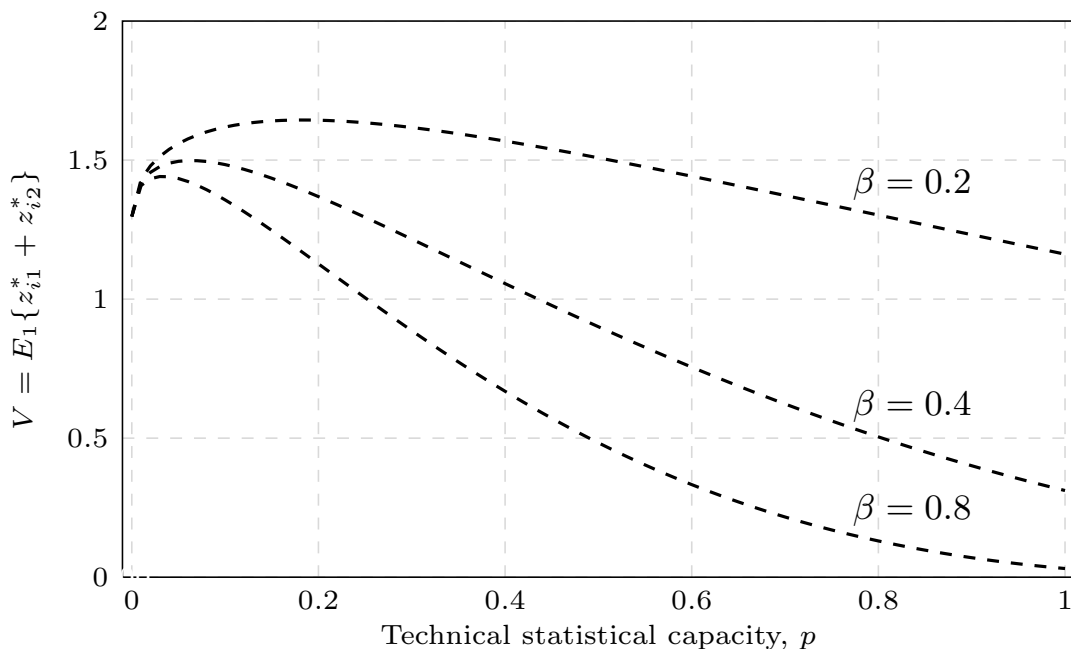
5.2 The Effect of Corruption

An increase in the bureaucracy's vulnerability to corruption strengthens the negative effect of statistical capacity on firm investment. As a result, the level of technical statistical capacity that maximizes comprehensive statistical capacity and informativeness, p^* , is a decreasing function of β (equation 30). Simulations suggest that a similar result holds for the level of statistical capacity that maximizes lifetime total output by the representative firm, p^{**} . Figure 4 illustrates that p^{**} shifts to the left as β increases. Consistent with this, we find that a rise in β has a negative effect on the two outcomes of interest:

PROPOSITION 6 *Suppose that the bureaucracy is sufficiently vulnerable to corruption so that condition (R1) holds. Then, (i) $C(p; \alpha, \beta)$ —and thus $I(C; \gamma, \theta)$ —are strictly decreasing functions of β ; (ii) $V = E_1 \{z_{i1}^* + z_{i2}^*\}$ is a strictly decreasing function of β .*

Proof. See Appendix A. ■

Figure 5: Effect of vulnerability to corruption on economic performance



For the simulation we have chosen the following parameter values: $\alpha = 0.7$, $\gamma = 0.5$.

There is a large literature on the channels by which corruption affects economic performance. Proposition 6 puts the spotlight on a novel one. An increase in the bureaucracy’s vulnerability to corruption reduces the informativeness of policy experiments (involving the implementation, evaluation, and adjustment of reforms) and hence slows down the learning process about “what works”. This is reflected in the economy’s growth rate:

PROPOSITION 7 *Suppose that the bureaucracy is sufficiently vulnerable to corruption so that condition (R1) holds. Then, the expected growth rate of modern-technology output,*

$$E_1 \left\{ \frac{y_{i2}^* - y_{i1}^*}{y_{i1}^*} \right\} = \frac{1}{1 - \gamma} \left\{ \hat{A}^l + I(C; \gamma, \theta) (\hat{A}^h - \hat{A}^l) \right\} - 1, \quad (32)$$

is a strictly decreasing function of β .

Proof. Follows immediately from Proposition 6. ■

So the consequences of an increase in corruption are not limited to a mere level effect; higher corruption also flattens the path of modern-technology output over time.

6 Summary and Conclusion

One approach to boost economic output is policy experimentation: tweak an existing economic policy, evaluate the consequences, and so discover “what works”. While economists have been interested in this approach for a long time, it has recently gained traction in the context of development policy. Clearly, how much policy makers can learn from a policy experiment—i.e., the degree of the experiment’s informativeness—must depend on the accuracy with which the economy is measured. In developing countries, accuracy tends to be low. As a result, efforts to help developing countries improve their statistical capacities are of great importance. In this paper, we demonstrate that efforts with a sole focus on technical aspects—here understood as the scale of data gathering—need not be unambiguously helpful. The reason is a type of “observer effect”: raising technical statistical capacity to improve the measurement of the economy changes what is being measured. Why? If confidentiality breaches are possible and control of corruption is weak, more extensive information gathering by the statistical office means that firms face a higher risk of being confronted with bribe demands. As a result, firms reduce their investments, thereby dampening the effect of policy changes. The relative strength of this harmful effect rises in the level of technical statistical capacity. At some point, it becomes the dominant force, implying that further improvements in technical statistical capacity make it harder—rather than easier—to discover “what works”.

Against this background, we argue that efforts with the aim of expanding data gathering in developing countries should not be uniform but adapted to local circumstances. Broadly speaking, the model suggests that such efforts be aligned with the quality of local institutions: in countries where control of corruption is weak and data confidentiality in doubt, the extent of data gathering—and hence the precision with which the economy is measured—should be more limited. We further argue that attempts to address insufficient statistical capacity would benefit from a comprehensive perspective that includes institutional aspects. The model suggests that taking measures to strengthen data confidentiality would dampen the observer effect and so permit the analysis of a less distorted economy, including the economy’s response to policy changes. Such measures could range from fortifying the statistical office’s independence from the bureaucracy to outright outsourcing of its tasks to an international body. More generally, we believe that this paper may contribute to a better understanding of the intricate ways through which corruption retards economic growth. By stifling private economic activity, and by limiting the optimal degree of measurement precision, corruption makes it harder for policy makers to discover how to tailor economic policies to local circumstances.

References

- Acemoglu, D., Antràs, P., and Helpman, E. (2007). Contracts and Technology Adoption. *American Economic Review*, 97(3):916–943.
- Aghion, P., Akcigit, U., Cagé, J., and Kerr, W. R. (2016). Taxation, corruption, and growth. *European Economic Review*, 86:24–51.
- Bai, J., Jayachandran, S., Malesky, E. J., and Olken, B. A. (2019). Firm Growth and Corruption: Empirical Evidence from Vietnam. *The Economic Journal*, 129(618):651–677.
- Berry, F., Iommi, M., Stanger, M., and Venter, L. (2018). The Status of GDP Compilation Practices in 189 Economies and the Relevance for Policy Analysis.
- Binswanger, J. and Oechslin, M. (2015). Disagreement and Learning about Reforms. *The Economic Journal*, 125(584):853–886.
- Binswanger, J. and Oechslin, M. (2019). Better Statistics, Better Economic Policies? *Unpublished manuscript*.
- Colonnelli, E. and Prem, M. (2017). Corruption and firms: evidence from randomized audits in brazil. *Unpublished Working Paper*.
- Croushore, D. (2011). Frontiers of Real-Time Data Analysis. *Journal of Economic Literature*, 49(1):72–100.
- Demirgüç-Kunt, A., Klapper, L., Singer, D., and Van Oudheusden, P. (2015). The Global Findex Database 2014: Measuring Financial Inclusion around the World. *World Bank Policy Research Working Paper 7255*.
- Devarajan, S. (2013). Africa’s Statistical Tragedy. *Review of Income and Wealth*, 59(Special Issue):9–15.
- Estrin, S., Korosteleva, J., and Mickiewicz, T. (2013). Which institutions encourage entrepreneurial growth aspirations? *Journal of Business Venturing*, 28(4):564–580.
- Fisman, R. and Svensson, J. (2007). Are corruption and taxation really harmful to growth? Firm level evidence. *Journal of Development Economics*, 83(1):63–75.
- Freund, C., Hallward-Driemeier, M., and Rijkers, B. (2016). Deals and delays: Firm-level evidence on corruption and policy implementation times. *World Bank Economic Review*, 30(2):354–382.

- GADM database of Global Administrative Areas (2012). Version 1. <http://www.gadm.org/>. Accessed May 2012.
- Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, 102(2):994–1028.
- Hodler, R. and Raschky, P. A. (2014). Regional Favoritism. *Quarterly Journal of Economics*, 129(2):995–1033.
- Jerven, M. (2013). *Poor Numbers*. Cornell University Press, Ithaca.
- Kiregyera, B. (2015). *The Emerging Data Revolution in Africa*. SUN MeDIA, Stellenbosch.
- Manski, C. F. (2015). Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern. *Journal of Economic Literature*, 53(3):631–653.
- Muralidharan, K. and Niehaus, P. (2017). Experimentation at Scale. *Journal of Economic Perspectives*, 31(4):103–124.
- National Oceanic and Atmospheric Administration (NOAA) (2014). Version 4 DMSP-OLS Nighttime Lights Time Series. Boulder, CO: National Geophysical Data Center. <https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>. Accessed January 2014.
- Olken, B. A. and Pande, R. (2012). Corruption in Developing Countries. *Annual Review of Economics*, 4(1):479–509.
- Orphanides, A. (2003). Monetary policy evaluation with noisy information. *Journal of Monetary Economics*, 50(3):605–631.
- Rodrik, D. (2010). Diagnostics before Prescription. *Journal of Economic Perspectives*, 24(3):33–44.
- Sandefur, J. and Glassman, A. (2015). The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics. *The Journal of Development Studies*, 51(2):116–132.
- Svensson, J. (2003). Who Must Pay Bribes and How Much? Evidence from a Cross-Section of Firms. *The Quarterly Journal of Economics*, 118(1):207–230.

Appendix A: Proofs

Propositions

Proof of Proposition 1. Obviously, the maximizer of $E_2\{[A_2(P_2)]^{1/(1-\alpha)}\}$ maximizes the expected second-period total output by the representative firm, $E_2\{z_{i2}^*(P_2)\}$. Together, equation (3) and the definition of $r(Z_1^n)$, stated in equation (8), imply

$$E_2\left\{[A_2(P_2)]^{\frac{1}{1-\alpha}}\right\} = \begin{cases} r(Z_1^n)(1 + \sqrt{\gamma})^{\frac{1}{1-\alpha}} + [1 - r(Z_1^n)](1 - \sqrt{\gamma})^{\frac{1}{1-\alpha}} & : P_2 = P_1 \\ 1 & : P_2 = 0 \\ r(Z_1^n)(1 - \sqrt{\gamma})^{\frac{1}{1-\alpha}} + [1 - r(Z_1^n)](1 + \sqrt{\gamma})^{\frac{1}{1-\alpha}} & : P_2 = -P_1 \end{cases} \quad (33)$$

Note that $E_2\{[A_2(P_1)]^{1/(1-\alpha)}\} \geq E_2\{[A_2(-P_1)]^{1/(1-\alpha)}\}$ if and only if $r(Z_1^n) \geq 1/2$. Moreover, because $[A_2]^{1/(1-\alpha)}$ is a strictly convex function of A_2 , we have $E_2\{[A_2(P_1)]^{1/(1-\alpha)}\} > 1$ if $r(Z_1^n) \geq 1/2$ and $E_2\{[A_2(-P_1)]^{1/(1-\alpha)}\} > 1$ if $r(Z_1^n) \leq 1/2$. As a result, the government will set $P_2 = P_1$ if $r(Z_1^n) \geq 1/2$ and $P_2 = -P_1$ if $r(Z_1^n) < 1/2$. ■

Proof of Proposition 3. The probability that the beneficial reform policy is implemented in period 2 is composed of the probability that $r(Z_1^p) \geq 1/2$ if $P_1 = S$ and the probability that $r(Z_1^p) < 1/2$ if $P_1 = -S$. Equation (22) implies (i) that $r(Z_1^p) \geq 1/2$ is equivalent to $Z_1^p \geq \hat{Z}_1^p$; and (ii) that $r(Z_1^p) < 1/2$ is equivalent to $Z_1^p < \hat{Z}_1^p$. Thus,

$$\begin{aligned} \Pr[P_2 = S] &= (1/2) \Pr\left[Z_1^p \geq \hat{Z}_1^p \mid P_1 = S\right] + (1/2) \Pr\left[Z_1^p < \hat{Z}_1^p \mid P_1 = -S\right] \\ &= (1/2) \left\{1 - \Pr\left[Z_1^p < \hat{Z}_1^p \mid P_1 = S\right] + \Pr\left[Z_1^p < \hat{Z}_1^p \mid P_1 = -S\right]\right\} \end{aligned} \quad (34)$$

Taking into account equations (3), (5), and (19), equation (34) can be rewritten as

$$\begin{aligned} \Pr[P_2 = S] &= (1/2) \left\{1 - \Pr\left[(1 + \sqrt{\gamma})[\alpha(1 - p\beta)]^{\alpha/(1-\alpha)} + \zeta_1^p < \hat{Z}_1^p\right]\right\} \\ &\quad + (1/2) \Pr\left[(1 - \sqrt{\gamma})[\alpha(1 - p\beta)]^{\alpha/(1-\alpha)} + \zeta_1^p < \hat{Z}_1^p\right], \end{aligned} \quad (35)$$

where \hat{Z}_1^p is given by equation (23). Rearranging terms yields

$$\begin{aligned} \Pr[P_2 = S] &= (1/2) \left\{1 - \Pr\left[\sqrt{p/\theta}(\zeta_1^p - \chi) < -\sqrt{\gamma/\theta} \cdot C\right]\right\} \\ &\quad + (1/2) \Pr\left[\sqrt{p/\theta}(\zeta_1^p - \chi) < \sqrt{\gamma/\theta} \cdot C\right], \end{aligned} \quad (36)$$

Since $\sqrt{p/\theta}(\zeta_1^p - \chi)$ follows a standard normal distribution, equation (36) is equivalent to

$$\Pr[P_2 = S] = (1/2) \left[1 - \Phi(-\sqrt{\gamma/\theta} \cdot C) + \Phi(\sqrt{\gamma/\theta} \cdot C)\right] \quad (37)$$

Because the standard normal distribution is symmetric around zero, equation (37) simplifies to the expression given in the proposition. ■

Proof of Proposition 4. As $\Phi\left(\sqrt{\gamma/\theta} \cdot C\right)$ is a strictly increasing function of $C(p; \alpha, \beta)$, the maximizer of C , p^* , is also the maximizer of $I = \Pr[P_2 = S]$. From equation (49), it follows that $\partial C(1; \alpha, \beta)/\partial p < 0$ if condition (R1) is satisfied. Thus, according to Lemma 2, the unique maximizer p^* is strictly less than 1 and hence must be pinned down by the first-order condition $\partial C(p^*; \alpha, \beta)/\partial p = 0$. Equation (49) suggests that this condition is equivalent to

$$\frac{1}{2} \frac{(1 - p^* \beta)}{\sqrt{p^*}} - \frac{\alpha}{1 - \alpha} \frac{\beta}{N} \sqrt{p^*} = 0. \quad (38)$$

Rearranging terms yields the expression stated in the proposition. ■

Proof of Proposition 5. As the government opts for a reform policy in period 1 ($P_1 = 1$), $V(p; \cdot)$ is given by equation (25). Therefore,

$$\begin{aligned} \frac{\partial V}{\partial p} &= \phi\left(\sqrt{\gamma/\theta} \cdot C\right) \frac{\partial C}{\partial p} \sqrt{\frac{\gamma}{\theta}} \left(\hat{A}^h - \hat{A}^l\right) [\alpha(1 - p\beta)]^{\alpha/(1-\alpha)} \\ &\quad - \alpha\beta \frac{\alpha}{1 - \alpha} \frac{1 + E_1 \left\{ A_2^{1/(1-\alpha)} \right\}}{[\alpha(1 - p\beta)]^{(1-2\alpha)/(1-\alpha)}} \end{aligned} \quad (39)$$

or, equivalently, by

$$\begin{aligned} \frac{\partial V}{\partial p} &= [\alpha(1 - p\beta)]^{(2\alpha-1)/(1-\alpha)} \\ &\quad \cdot \left\{ \phi\left(\sqrt{\gamma/\theta} \cdot C\right) \frac{\partial C}{\partial p} \sqrt{\frac{\gamma}{\theta}} \left(\hat{A}^h - \hat{A}^l\right) [\alpha(1 - p\beta)] - \alpha\beta \frac{\alpha}{1 - \alpha} \left[1 + E_1 \left\{ A_2^{1/(1-\alpha)} \right\}\right] \right\}. \end{aligned} \quad (40)$$

First suppose that $\alpha \leq 1/2$ and consider equation (39). As p rises from 0 to the maximizer of $C(p; \alpha, \beta)$, p^* , the first line of equation (39) decreases monotonically from infinity to 0 (see the properties of $C(p; \alpha, \beta)$ stated in Lemma 2); for values of $p \in (p^*, 1]$, the first line is strictly negative. The second line of equation (39) increases monotonically (in absolute terms) as p rises from 0 to p^* (see equation 26 and Proposition 3), and remains strictly negative when p exceeds p^* (and rises towards 1). Together, the behavior of the first and the second line implies that $\partial V/\partial p$ —as p rises from 0 to p^* —falls monotonically from infinity to a value that is strictly less than 0; for values of $p \in (p^*, 1]$, $\partial V/\partial p$ remains strictly negative. As a result, on $[0, 1]$, there exists a unique $p^{**} < p^*$ such that $\partial V(p^{**}; \cdot)/\partial p = 0$. From this, it immediately follows that p^{**} is the unique maximizer of V . Now suppose $\alpha > 1/2$ and consider equation (40). A similar chain of arguments implies that, again, on $[0, 1]$, there exists a unique $p^{**} < p^*$ such that $\partial V(p^{**}; \cdot)/\partial p = 0$. As in the case $\alpha \leq 1/2$, p^{**} is the unique maximizer of V . ■

Proof of Proposition 6. To prove (i), first assume that $dp^{**}/d\beta > 0$. In this case, we consider $\partial V/\partial p|_{p=p^{**}} = 0$, the first-order condition that implicitly pins down p^{**} . Using

equations (39) and (49), and rearranging terms, we obtain

$$\frac{\hat{A}^h - \hat{A}^l}{2} \frac{\phi\left(\sqrt{\gamma/\theta} \cdot C\right) \sqrt{\gamma/\theta} \cdot C}{1 + E_1\{A_2^{1/(1-\alpha)}\}} \left(\frac{1}{p^{**}} - \frac{1+\alpha}{1-\alpha}\beta\right) = \frac{\alpha}{1-\alpha}\beta. \quad (41)$$

Expressions for C and $E_1\{A_2^{1/(1-\alpha)}\}$ are given by equations (29) and (26), respectively. Taking these into account, we can transform equation (41) into

$$U(C) \frac{\hat{A}^h - \hat{A}^l}{2} \sqrt{\frac{\gamma}{\theta}} [\alpha(1 - p^{**}\beta)]^{\alpha/(1-\alpha)} \left(\frac{1}{\sqrt{p^{**}}} - \frac{1+\alpha}{1-\alpha}\beta\sqrt{p^{**}}\right) = \frac{\alpha}{1-\alpha}\beta, \quad (42)$$

where

$$U(C) = \frac{\phi\left(\sqrt{\gamma/\theta} \cdot C\right)}{1 + \hat{A}^l + \Phi\left(\sqrt{\gamma/\theta} \cdot C\right) \left(\hat{A}^h - \hat{A}^l\right)}. \quad (43)$$

Now consider the impact of a rise in β on equation (42), holding constant C . Obviously, the right-hand side of the equation increases; at the same time, since we assume $dp^{**}/d\beta > 0$, the left-hand side decreases. As a result, $U(C)$ must rise. Since $C > 0$ and Φ and ϕ denote, respectively, the distribution function and the probability density function of the standard normal distribution, it follows that C must fall to restore the equality of the two sides of equation (42). So, if $dp^{**}/d\beta > 0$, we must conclude that $C(p^{**}; \alpha, \beta)$ is a strictly decreasing function of β . Now suppose that $dp^{**}/d\beta \leq 0$ and consider

$$\frac{dC(p^{**}; \alpha, \beta)}{d\beta} = \frac{\partial C(p; \alpha, \beta)}{\partial p} \Big|_{p=p^{**}} \frac{dp^{**}}{d\beta} + \frac{\partial C(p; \alpha, \beta)}{\partial \beta} \Big|_{p=p^{**}}. \quad (44)$$

Observing equation (29), we can infer that $\partial C(p; \alpha, \beta)/\partial \beta|_{n=n^{**}} < 0$. Because $p^{**} < p^*$, and since $C(p; \alpha, \beta)$ is a strictly increasing function of p on $[0, p^*]$, we have $\partial C(p; \alpha, \beta)/\partial p|_{p=p^{**}} > 0$. Finally, because $dp^{**}/d\beta \leq 0$ by assumption, it follows that $dC(p^{**}; \alpha, \beta)/d\beta < 0$. So, once again, we must conclude that $C(p^{**}; \alpha, \beta)$ is a strictly decreasing function of β .

To prove (ii), note that equations (25) and (26) and Proposition 3 imply that in equilibrium

$$V = \left[1 + \hat{A}^l + \Phi\left(\sqrt{\gamma/\theta} \cdot C(p^{**}; \alpha, \beta)\right) \left(\hat{A}^h - \hat{A}^l\right)\right] [\alpha(1 - p^{**}\beta)]^{\alpha/(1-\alpha)} + 2\chi. \quad (45)$$

Obviously, $\partial V/\partial p|_{p=p^{**}} = 0$. Therefore,

$$\begin{aligned} \frac{dV}{d\beta} \Big|_{p=p^{**}} &= \phi\left(\sqrt{\gamma/\theta} \cdot C\right) \sqrt{\gamma/\theta} \frac{\partial C}{\partial \beta} \left(\hat{A}^h - \hat{A}^l\right) [\alpha(1 - p^{**}\beta)]^{\alpha/(1-\alpha)} \\ &\quad - \left[1 + \hat{A}^l + \Phi\left(\sqrt{\gamma/\theta} \cdot C\right) \left(\hat{A}^h - \hat{A}^l\right)\right] \frac{\alpha}{1-\alpha} [\alpha(1 - p^{**}\beta)]^{\alpha/(1-\alpha)-1} \alpha p^{**}. \end{aligned} \quad (46)$$

Considering the definition of C given in equation (29), it follows immediately that $\partial C/\partial \beta < 0$. Given this, we conclude that $dV/d\beta|_{p=p^{**}} < 0$. ■

Lemmas

Proof of Lemma 1. In period 2, the government will either keep the first-period reform policy in place or switch to the alternative reform policy (equation 18). Given the definition of A_2 in equation (3), and that of $\Pr[P_2 = S]$ in the lemma, the agents expect $A_2 = 1 + \sqrt{\gamma}$ with probability $\Pr[P_2 = S]$ and $A_2 = 1 - \sqrt{\gamma}$ with probability $1 - \Pr[P_2 = S]$. Thus,

$$E_1 \left\{ [A_2(P_2)]^{1/(1-\alpha)} \right\} = \Pr[P_2 = S](1 + \sqrt{\gamma})^{1/(1-\alpha)} + \{1 - \Pr[P_2 = S]\}(1 - \sqrt{\gamma})^{1/(1-\alpha)}, \quad (47)$$

an expression that can be rearranged to give the expression in the lemma.

To see that $E_1 \{ [A_2(P_2)]^{1/(1-\alpha)} \} > 1$, note (i) that $[A_2]^{1/(1-\alpha)}$ is a strictly convex function of A_2 , implying that $(1/2)(1 + \sqrt{\gamma})^{1/(1-\alpha)} + (1/2)(1 - \sqrt{\gamma})^{1/(1-\alpha)} > 1$; and (ii) that $\Pr[P_2 = S]$ is strictly greater than $1/2$ since the observation of Z_1^p permits learning. ■

Proof of Lemma 2. The partial derivative of C with respect to p is given by

$$\frac{\partial C}{\partial p} = \sqrt{\frac{\gamma}{\theta}} \left[\frac{1}{2} \frac{[\alpha(1-p\beta)]^{\alpha/(1-\alpha)}}{\sqrt{p}} - \alpha\beta \frac{\alpha}{1-\alpha} \frac{\sqrt{p}}{[\alpha(1-p\beta)]^{(1-2\alpha)/(1-\alpha)}} \right] \quad (48)$$

or, equivalently, by

$$\frac{\partial C}{\partial p} = \alpha^{\alpha/(1-\alpha)} (1-p\beta)^{(2\alpha-1)/(1-\alpha)} \sqrt{\frac{\gamma}{\theta}} \left[\frac{1}{2} \frac{1-p\beta}{\sqrt{p}} - \frac{\alpha}{1-\alpha} \beta \sqrt{p} \right]. \quad (49)$$

First suppose that $\alpha \leq 1/2$. Then, equation (48) implies that $\lim_{p \rightarrow 0} \partial C / \partial p = \infty$ and that $\partial C / \partial p$ is a strictly decreasing function of p on $[0, 1]$. We thus conclude that C is strictly concave on $[0, 1]$ and has a unique maximizer, p^* , that is strictly greater than 0. Now suppose $\alpha > 1/2$. Then, equation (49) implies that $\lim_{p \rightarrow 0} \partial C / \partial p = \infty$ and that—as p increases from 0— $\partial C / \partial p$ is a strictly decreasing function of p as long as $\partial C / \partial p \geq 0$. Equation (49) further suggests that $\partial C / \partial p$ can have at most one root on $[0, 1]$. Thus, again, C has a unique maximizer $p^* > 0$; moreover, it is strictly concave on $[0, p^*]$. ■

Appendix B: Endogenous Bribe Demands

Assumptions and approach. In the baseline model, we maintain the assumption that firms belonging to the random sample may be subject to bribe demands, while firms outside the sample are left alone for sure. A simple extension of the baseline model allows us to obtain this pattern as an endogenous equilibrium outcome. To simplify matters, suppose that confidentiality is never adhered to: $\pi = 1$, so that $\beta = \pi\hat{\beta} = \hat{\beta}$. Now assume that, immediately after the statistical office has drawn the random sample, the corrupt official (i) learns the

identity of the firms in the sample; (ii) must decide which firms, if any, to approach for initiation of the bribe extraction process; (iii) incurs a cost $\psi > 0$ for each firm approached. After the statistical office has collected the data, the corrupt official can present the firms for which the extraction process has been initiated with bribe demands; when presenting the demands, thanks to the confidentiality breach, the corrupt official knows the output of those firms that belong to the sample, but lacks this knowledge for firms that are not in the sample. In what follows, we demonstrate that the equilibrium characterized in the main text is consistent with the official exclusively approaching sampled firms if the cost of approaching firms, ψ , takes an appropriate value.

Notation and preliminaries. To show that the differential treatment of sampled and non-sampled firms can be an equilibrium outcome, we introduce additional notation. F and f will stand for the distribution function and probability density function of ζ_{kt} , respectively. We denote by \mathcal{R}_t the set of firms sampled in period t and by d_{kt} the bribe demand presented to firm k in period t . We use $e(d_{kt})$ to denote the bribe received by the corrupt official from firm k if the firm has been presented with a demand of size d_{kt} . Taking into account that firm k can fend off bribe demands at a cost of $\hat{\beta}z_{kt}$ (see Subsection 3.3), we obtain

$$e(d_{kt}) = d_{kt} \cdot \mathbb{1}[d_{kt} \leq \hat{\beta}z_{kt}], \quad (50)$$

where $\mathbb{1}[d_{kt} \leq \hat{\beta}z_{kt}]$ is a dummy variable that takes a value of 1 if the condition in brackets is satisfied (and a value of 0 otherwise). Taking into account equations (1) and (13), and the fact that y_{kt}^* depends on whether $P_t = S$ or $P_t \neq S$, we obtain

$$e(d_{kt}|P_t \neq S) = d_{kt} \cdot \mathbb{1}\left[d_{kt}/\hat{\beta} - y_{kt}^*|_{P_t \neq S} \leq \zeta_{kt}\right] \quad (51)$$

and

$$e(d_{kt}|P_t = S) = d_{kt} \cdot \mathbb{1}\left[d_{kt}/\hat{\beta} - y_{kt}^*|_{P_t = S} \leq \zeta_{kt}\right], \quad (52)$$

where we will work with the definitions $\zeta_t^h(d_{kt}) \equiv d_{kt}/\hat{\beta} - y_{kt}^*|_{P_t \neq S}$ and $\zeta_t^l(d_{kt}) \equiv d_{kt}/\hat{\beta} - y_{kt}^*|_{P_t = S}$ in what follows. We use the superscripts “ h ” and “ l ” to indicate that the threshold is at a high level if $P_t \neq S$ (because of low modern-technology output) and at a low level if $P_t = S$ (because of high modern-technology output). Relying on this notation, we now can express the expected bribe payment by firm k as a function of the bribe demand:

$$E_t^{co}\{e(d_{kt})\} = (1 - Q_t)E_t^{co}\{e(d_{kt})|P_t \neq S\} + Q_tE_t^{co}\{e(d_{kt})|P_t = S\}, \quad (53)$$

where $E_t^{co}\{\cdot\}$ refers to the expectation formed in period t at the moment the corrupt official (hence the superscript “ co ”) must decide whether or not to approach firm k for initiation of

the bribe extraction process and Q_t denotes the corresponding probability that $P_t = S$. Note that $Q_1 = 1/2$, while Q_2 is either equal to $r(Z_1^p)$ if this magnitude is at least $1/2$ or equal to $1 - r(Z_1^p)$ otherwise (in the latter case, the government switches to the alternative reform policy in period 2). Importantly, this implies that $Q_2 \in [1/2, 1)$. Finally, making use of equations (51) and (52), equation (53) can be rewritten as

$$\begin{aligned} E_t^{co} \{e(d_{kt})\} &= (1 - Q_t) \int_0^\infty d_{kt} \mathbb{1} \left[\zeta_t^h(d_{kt}) \leq \zeta_{kt} \right] f(\zeta_{kt}) d\zeta_{kt} \\ &+ Q_t \int_0^\infty d_{kt} \mathbb{1} \left[\zeta_t^l(d_{kt}) \leq \zeta_{kt} \right] f(\zeta_{kt}) d\zeta_{kt}. \end{aligned} \quad (54)$$

From equation (54), one can infer that—in expectations—the bribe extracted from a random firm $i \in \mathcal{R}_t$ is strictly greater than the bribe that could be extracted from a random firm $j \notin \mathcal{R}_t$. In the case of firm i , the corrupt official observes total firm output and hence can always adjust d_{it} such that—given the realization of ζ_{it} and the relationship between P_t and S —the maximum bribe is extracted. In the case of firm j , output is not observed, implying that the official has to ask for a uniform bribe. However, doing so means that—given the realization of ζ_{jt} and the relationship between P_t and S —the bribe received is either 0 or strictly less than the maximum possible.¹⁵ So, for firm j , the integrals in equation (54) take a strictly lower value than do the corresponding integrals for firm i .

Minimum expected bribe: random firm $i \in \mathcal{R}_t$. To find a condition that ensures that the corrupt official will always find it optimal to approach firms belonging to the statistical office's random sample, we have to derive the minimum value of $E_t^{co} \{e(d_{it})\}$, where $d_{it} = \hat{\beta} z_{it}$ since $i \in \mathcal{R}_t$. The minimum of $E_t^{co} \{e(d_{it})\}$ must arise in period 1 because $Q_1 = 1/2$, which is the lowest possible value probability Q_t can take (see the discussion above). Using $d_{i1} = \hat{\beta} z_{i1}$ and $Q_1 = 1/2$ in equation (54), we obtain

$$\begin{aligned} E_1^{co} \left\{ e(\hat{\beta} z_{i1}) \right\} &= \frac{1}{2} \int_0^\infty \hat{\beta} \left(y_{i1}^* |_{P_1 \neq S} + \zeta_{i1} \right) f(\zeta_{i1}) \zeta_{i1} \\ &+ \frac{1}{2} \int_0^\infty \hat{\beta} \left(y_{i1}^* |_{P_1 = S} + \zeta_{i1} \right) f(\zeta_{i1}) \zeta_{i1}, \end{aligned} \quad (55)$$

where $y_{i1}^* |_{P_1 \neq S}$ can be found by combining equations (3) and (19). Simplifying yields

$$E_1^{co} \left\{ e(\hat{\beta} z_{i1}) \right\} = \hat{\beta} \left\{ [\alpha (1 - p^{**} \beta)]^{\alpha/(1-\alpha)} + \chi \right\}. \quad (56)$$

If the expression in equation (56) is greater than ψ , the corrupt official will always find it optimal to approach all firms $i \in \mathcal{R}_t$ in order to ask them for bribes.

¹⁵Of course, the realization of ζ_{jt} may be such that it exactly hits the relevant threshold. However, since ζ_{jt} has a continuous distribution, the chance that it takes this particular threshold is 0.

Maximum expected bribe: random firm $j \notin \mathcal{R}_t$. When the corrupt official considers approaching a random firm $j \notin \mathcal{R}_t$, he knows that he will not be in a position to adjust his bribe demand to the firm’s output level. Instead, if he approached the firm, he would ask for a uniform bribe that corresponds to the maximizer of equation (54). In what follows, we denote this maximizer by \bar{d}_{jt} . To find a condition that ensures that the corrupt official will never find it optimal to approach firms outside the statistical office’s random sample, we have to find the highest value that $E_t^{co} \{e(\bar{d}_{jt})\}$ can attain. The highest value arises when Q_t —the probability that $P_t = S$ —equals 1. Obviously, this can only happen in period 2. Taking all these considerations into account, it follows from equation (54) that

$$E_2^{co} \{e(\bar{d}_{j2})\} = \bar{d}_{j2} \left[1 - F \left(\zeta_2^l(\bar{d}_{j2}) \right) \right] \quad (57)$$

is the highest bribe payment the corrupt official could expect if he approached an arbitrary non-sampled firm. Using the definition of $\zeta_t^l(d_{jt})$, and the explicit expression for $y_{j2}^*|_{P_2=S}$ inferable from Subsection 4.2, equation (57) can be turned into

$$E_2^{co} \{e(\bar{d}_{j2})\} = \hat{\beta} \left\{ (1 + \sqrt{\gamma})^{1/(1-\alpha)} [\alpha (1 - p^{**}\beta)]^{\alpha/(1-\alpha)} + \zeta_2^l(\bar{d}_{j2}) \right\} \cdot \left[1 - F \left(\zeta_2^l(\bar{d}_{j2}) \right) \right]. \quad (58)$$

If the expression given in equation (58) is strictly less than ψ , the corrupt official will never find it optimal to approach a firm $j \notin \mathcal{R}_t$ in order to ask for bribes.

Differential treatment of sampled and non-sampled firms. The discussion in the previous two paragraphs immediately implies that, if the condition

$$E_1^{co} \left\{ e(\hat{\beta}z_{i1}) \right\} > \psi > E_2^{co} \{e(\bar{d}_{j2})\} \quad (59)$$

holds, it is optimal for the corrupt official to approach all firms $i \in \mathcal{R}_t$ but to leave alone all firms $j \notin \mathcal{R}_t$ in period $t \in \{1, 2\}$. The above ranking of expectations is guaranteed provided that χ is sufficiently large, an assumption we can make without further ado since χ does not affect any of the endogenous variables. Given this, it follows that for “intermediate” values of ψ the corrupt official behaves exactly as is assumed in the baseline setup.

Appendix C: Misreporting

Assumptions. In the model, we maintain the assumption that sampled firms have no choice but to report output truthfully. In what follows, we demonstrate that the model can be

extended to allow for output misreporting without undermining its key implications. As in Appendix B, for simplicity, we assume that confidentiality is never adhered to (i.e., $\pi = 1$). Moreover, we again use \mathcal{R}_t to denote the set of firms sampled in period t .

Now assume that, (i) at the time the statistical office draws its random sample, the corrupt official commits to a “two-part bribe tariff” that will be applied to sampled firms: $d_{kt} = d_t^0 + d_t^1 z_{kt}^r$, where (d_t^0, d_t^1) is chosen by the official and z_{kt}^r refers to output as *reported* by firm $k \in \mathcal{R}_t$; (ii) once the ζ_{it} s are determined, each sampled firm simultaneously decides on z_{kt}^r and on whether to accept or fend off the bribe demand, considering that with probability $\pi^d > 0$ misreporting carries a fine of $(z_{kt} - z_{kt}^r)^2$ and that fending the demand off comes at a cost of $\hat{\beta} z_{kt}$ (as in the baseline setup). Assume further that (iii) the corrupt official is anxious to avoid that even only a single bribe demand is fended off (as this might draw unwelcome attention to the official’s machinations, something the official wants to avoid at all costs); in formal terms, suppose that, for the official, a bribe demand that is fended off carries an infinitively large cost. Assumption (iii) is not necessary in the baseline setup, but it is perfectly compatible with it. Finally, assume that (iv) the statistical office understands (i) to (iii).

Equilibrium (mis-)reporting. When determining z_{kt}^r and its response to the bribe demand, firm $k \in \mathcal{R}_t$ minimizes (what we call) its expected “cost of reporting”, thereby taking z_{kt} , d_t^0 , and d_t^1 as given. The cost of reporting consists of the punishment for misreporting (if any) and bribe-related expenses (payment or cost of fending the demand off). Since accepting the bribe demand or fending it off is a binary choice, firm k proceeds as follows. For each of the two options, it identifies the level of z_{kt}^r that minimizes the expected cost of reporting. It is straightforward to identify the optimal choice of z_{kt}^r if the demand is fended off. Since the cost of fending it off is $\hat{\beta} z_{kt}$, and therefore independent of z_{kt}^r , it follows immediately that sticking to the truth (i.e., $z_{kt}^r = z_{kt}$) minimizes the expected cost of reporting. So, if firm k chooses to fend off the bribe demand, the expected cost of reporting is simply $\hat{\beta} z_{kt}$.

Otherwise, if the bribe demand is accepted, firm k has to solve the minimization problem

$$\min_{z_{kt}^r} \left\{ \pi^d (z_{kt} - z_{kt}^r)^2 + (d_t^0 + d_t^1 z_{kt}^r) \right\}. \quad (60)$$

The objective function in problem (60) is strictly convex. Its solution is

$$z_{kt}^r = z_{kt} - \frac{1}{2} \frac{d_t^1}{\pi^d}, \quad (61)$$

an expression we assume to be non-negative (which is verified below). Given this, a simple calculation yields that accepting the bribe demand means an expected cost of reporting of $d_t^0 + d_t^1 z_{kt} - [(d_t^1/2)^2/\pi^d]$. In the baseline setup, we assume that firms prefer accepting the

bribe demand to fending it off if the costs are the same. Here we stick to this assumption. Firm k therefore prefers paying the bribe if and only if

$$d_t^0 + d_t^1 z_{kt} - [(d_t^1/2)^2/\pi^d] \leq \hat{\beta} z_{kt}. \quad (62)$$

Now consider the corrupt official. Anticipating how firms will respond to bribe demands, the official chooses (d_t^0, d_t^1) so as to maximize expected bribe payments—subject to the constraint that the resulting bribe tariff does not drive any firm into fending off a demand. From the discussion in the preceding paragraph we know that an arbitrary firm $k \in \mathcal{R}_t$ accepts the bribe demand if and only if condition (62) is satisfied. It is easy to see that the two sides of condition (62) are identical if $d_t^0 = [(d_t^1/2)^2/\pi^d]$ and $d_t^1 = \hat{\beta}$. In graphical terms, both sides are linear functions of z_{kt} with a slope of $\hat{\beta}$. Therefore, condition (62) holds with equality for any possible value of $z_{kt} \geq \underline{z}_{kt} > 0$, where \underline{z}_{kt} is the lower bound on z_{kt} in period t . In fact, from the perspective of the corrupt official, this (d_t^0, d_t^1) -combination is the optimal one. As a graphical analysis immediately shows, other combinations either imply that there are z_{kt} -ranges for which condition (62) is violated (positive chance that a bribe demand is fended off) or that condition (62) is non-binding for any possible value of z_{kt} (bribe extracted always lower than under the optimal combination). So, in equilibrium,

$$(d_t^0, d_t^1) = \left((\hat{\beta}/2)^2/\pi^d, \hat{\beta} \right), \quad (63)$$

which implies

$$d_{kt} = \frac{(\hat{\beta}/2)^2}{\pi^d} + \hat{\beta} \cdot z_{kt}^r \quad \text{and} \quad z_{kt}^r = z_{kt} - \frac{1}{2} \frac{\hat{\beta}}{\pi^d}. \quad (64)$$

As a result, for firm $k \in \mathcal{R}_t$, the equilibrium expected cost of reporting (which now consists of the bribe payment plus the expected fine due to misreporting) is simply given by $\hat{\beta} z_{kt}$, i.e., by the same magnitude as in the baseline setup (with $\pi = 1$). At this point, also note that the statistical office's estimate of average total firm output, Z_t^p , is as informative as it is in the baseline setup: as the statistical office understands the official's and the firms' decision problems, it can simply infer the actual output levels from the reported magnitudes via the corresponding expression in equation (64).

The fact that the expected cost of reporting is unchanged implies that firms' investment incentives are unchanged, too. In particular, firms still solve maximization problem (11), where $\beta = \hat{\beta}$ since $\pi = 1$. So y_{it}^* is again given by equation (13). Because of this, and since there is no difference in the properties of Z_t^p , the government's incentives do not change either. So the rest of the analysis is as in the baseline setup, too. Thus, in the end, we obtain the same equilibrium level of technical statistical capacity, p^{**} .

To close the formal discussion, we return to equation (61), in the context of which we assumed z_{kt}^r to be non-negative. This requires $\min\{z_{k1}, z_{k2}\} \geq \hat{\beta}/(2\pi^d)$. As the minimum value of ζ_{kt} is 0, z_{kt} is given by the minimum value of y_{it}^* . The lowest possible modern-technology output materializes in period 2, in the case of $P_2 \neq S$. Then, $y_{k2}^* = \hat{A}^l \left[\alpha \left(1 - p^{**} \hat{\beta} \right) \right]^{\alpha/(1-\alpha)}$, where \hat{A}^l is given by equation (27). So $z_{kt}^r \geq 0$ requires

$$\hat{A}^l \left[\alpha \left(1 - p^{**} \hat{\beta} \right) \right]^{\alpha/(1-\alpha)} \geq \frac{1}{2} \frac{\hat{\beta}}{\pi^d}, \quad (65)$$

a condition that holds for sufficiently small values of $\hat{\beta}$.

We conclude that misreporting leaves the substance of the analysis in the main text unchanged. Still, there are two notable differences to the equilibrium in the baseline setup. First, the possibility of misreporting reduces the amount of bribes the corrupt official can extract. The difference is equal to the total amount of fines for misreporting. Second, since the corrupt official applies a two-part tariff, bribes as a share of reported output, d_{kt}/z_{kt}^r , decreases in z_{kt}^r . This is consistent with recent evidence presented by (Bai et al., 2019), who show that in Vietnam smaller firms pay higher bribes as a percentage of income.