Individuals and Organizations as Sources of State Effectiveness

By Michael Carlos Best and Jonas Hjort and David Szakonyi*

Bureaucrats implement policy. How important are they for a state's productivity? And do the tradeoffs between policies depend on their effectiveness? Using data on 16 million public purchases in Russia, we show that 39 percent of the variation in prices paid for narrowly defined items is due to the individual bureaucrats and organizations who manage procurement. Low-price buyers also display higher spending quality. Theory suggests that such differences in effectiveness can be pivotal for policy design. To illustrate, we show that a common one—bid preferences for domestic suppliers—substantially improves procurement performance, but only when implemented by ineffective bureaucrats.
JEL: H11, D73, O12

A successful state is the foundation economic development is built on (Besley and Persson, 2009; Page and Pande, 2018). States delegate policy implementation to their middle management tier, the bureaucracy. Historically, the dominant view in social science was that states could and should strive for a Weberian bureaucracy—"machines" mindlessly translating policy into output, ensuring uniform provision of public services (Weber, 1921). In reality, the skills, organizational capacity, and priorities of bureaucrats differ. But by how much? And what are the implications for policy design?

This paper aims to advance our understanding of the state's production func-

* Best: Columbia, New York, michael.best@columbia.edu. Hjort: UCL, London, j.hjort@ucl.ac.uk. Szakonyi: George Washington, Washington, D.C., dszakonyi@gwu.edu. We are grateful to editor Esther Duflo and two anonymous referees for comments that substantially improved the paper. We also thank Ben Olken for his very helpful discussion of the paper at the AEA meetings, and John M. Abowd, Daron Acemoglu, Hunt Allcott, Tim Besley, Nick Bloom, Richard Blundell, Raj Chetty, Allan Collard-Wexler, Francesco Decarolis, Andres Drenik, Pascaline Dupas, Ben Faber, Fred Finan, Matthew Gentzkow, Josh Gottlieb, Caroline Hoxby, Amit Khandelwal, Brad Larsen, David Margolis, Torsten Persson, Andrea Prat, Imran Rasul, Jimmy Roberts, Jonah Rockoff, Tobias Salz, Orie Shelef, Andrei Shleifer, David Silver, Eric Verhoogen, Guo Xu, and Danny Yagan for valuable comments; Andrei Yakovlev for institutional guidance; Andrey Druzhaev, Len Goff, Vinayak Iyer, and Georgiy Syunyaev for outstanding research assistance; and seminar participants at the AEA meetings, U.C. Berkeley, University of British Columbia, Central European University, Columbia, University of Copenhagen, Duke, the Econometric Society, Harvard Business School, Higher School of Economics, McGill, University of Michigan, the NBER Development Economics Program meeting, NEUDC, NYU, Northwestern, University of Oslo, Paris School of Economics, Santa Clara, Stanford, UCL, UCLA, University of Toronto, University of Virginia, and the World Bank for comments. Best thanks the Stanford Institute for Economic Policy Research, where he was a postdoctoral fellow while much of the work on the paper was done. The study has been funded within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and by the Russian Academic Excellence Project "5-100". All remaining errors are ours alone.

tion, an object that remains almost entirely unknown.[1] Our goals are two-fold. First, to quantify the importance of the bureaucracy for the productivity of the state. Second, to explore how the tradeoffs between different policies depend on the effectiveness of the bureaucracy in charge of implementation. The second goal is of particular importance in the public sector, where policy design may be relatively malleable compared to modifying hiring, training, and incentive practices to directly improve bureaucratic effectiveness. Both goals are challenging, as bureaucracies produce a wide array of outputs that are difficult to measure. However, one task—the procurement of off-the-shelf goods—is performed throughout the state enterprise, and has a well-defined and quantifiable primary output: prices paid.

We use a simple conceptual framework of procurement with endogenous supplier entry to guide our analysis of administrative data covering the universe of public procurement in Russia. With an empirical specification derived from the model, we estimate that 39 percent of the variation in performance—quality-adjusted prices paid—is attributable to the bureaucrats who manage procurement, roughly half to individual procurement officers and half to the end-user public organizations. The procurers also explain 24 percent of the variation in spending quality, and price- and quality-effectiveness are positively correlated. Differences in effectiveness of such magnitude have far-reaching implications for policy design. To illustrate, we examine the introduction of a bid preference regime common throughout the world. Under Russia's bid preferences, contract-winners offering goods manufactured abroad are paid 85 percent of their bid. Consistent with our model's predictions, we find that preferences substantially reduce costs and increase competitiveness, but only when the policy is implemented by *ineffective* bureaucrats.

Public procurement in Russia is an ideal setting to study micro-level state effectiveness. First, procurement makes up roughly 8 percent of worldwide GDP (Schapper, Malta and Gilbert, 2009). Second, for purchases of items that are precisely defined ("off-the-shelf" goods), procurers' mandate is simply to pay the lowest possible price while following the government's policy rules (see also Bandiera, Prat and Valletti, 2009; Ferraz, Finan and Szerman, 2015).[2] This makes performance measurable and comparable across the entire state enterprise. Third, Russia's massive and diverse bureaucracy spans a wide range of state effectiveness. Finally, the labor market for Russian procurement officers is decentralized and the resulting private-sector-like churn makes it possible to identify individuals' and their employers' effectiveness.

In our stylized model, bureaucratic effectiveness affects procurement outcomes in two ways. First, ineffective bureaucracies impose costs (e.g. unusual product specifications) that raise the cost to suppliers of fulfilling the contract. Second,

[1] This is despite a growing literature on front-line public sector workers (see e.g. Finan, Olken and Pande, 2017, for an overview).

[2] Russia spends over half of its total public procurement budget on such goods.

ineffective bureaucracies impose higher participation costs (e.g. required deposits, or bribes to enter the auction) on bidders. As a result, they attract fewer participants and pay higher quality-adjusted prices.

To compare the performance of bureaucrats (procurement officers) and organizations (e.g. ministries, schools or hospitals) across the country empirically, we need to ensure that they are performing the same task—buying the same type and quality of good. To do this, we adapt tools from machine learning to develop a methodology that uses the text of procurement contracts to classify purchases into homogeneous bins.[3] We also confirm that our results are very similar in a subsample of goods that are by nature homogeneous—pharmaceuticals—and alongside price-effectiveness consider spending quality outcomes (such as delays, contract renegotiation, and cost over-runs). For identification we exploit the fact that many organizations are observed working with multiple bureaucrats and vice versa. This provides us with thousands of quasi-experiments that capture the causal impact of individual bureaucrats and organizations on prices paid under weak assumptions on bureaucrat–organization matching. Event studies reveal large and sharp decreases in quality-adjusted prices paid when organizations switch to more effective bureaucrats, and vice versa, supporting a causal interpretation of these effects.[4]

To aggregate the impacts of individual bureaucrats and organizations on prices paid into an estimate of the share of the total variation explained by the bureaucratic apparatus as a whole, we extend the variance decomposition approach pioneered by Abowd, Kramarz and Margolis (1999); Abowd, Creecy and Kramarz (2002) (hereafter AKM) in two ways. First, we correct the fixed-effect estimates for sampling error using split-sample methods (Finkelstein, Gentzkow and Williams, 2016; Silver, 2016), and by extending shrinkage methods (Kane and Staiger, 2008; Chetty, Friedman and Rockoff, 2014*a*) to a two-dimensional context to explicitly account for the covariance between the estimation error in the bureaucrat and the organization effects (Andrews et al., 2008).[5] Second, we show how to estimate lower bounds on the variation explained by bureaucrats and organizations in a setting—like ours—where bureaucrat switching does not link *all* organizations and how the combined productivity effect of bureaucrats and organizations can nevertheless be identified.

[3]Our methodology ensures that within-category quality differences are minimal, while maintaining generality by not restricting to very specific types of goods. In foregoing conventional methods for categorizing comparable goods and instead using text analysis, we follow Hoberg and Phillips (2016). They classify *firm* similarity based on the goods produced.

[4]Importantly, our estimates can be interpreted causally even if bureaucrats sort across organizations based on the effectiveness of the bureaucrat and/or the organization. Instead, the assumptions needed for causal interpretation are that they do not sort based on unmodelled *match* effects, and that drift in effectiveness and switches are uncorrelated. The event studies provide compelling support for these assumptions, as does a battery of additional tests. Studies of the wages of workers and firms in the private sector tend to find the same (see Card et al. (2018); Bloom et al. (2019) for overviews of the literature).

[5]To our knowledge, two-dimensional shrinkage estimators like the ones we develop have not been used before.

We find that the bureaucracy jointly accounts for 39 percent of the variation in quality-adjusted prices paid, of which individuals and organizations account for roughly equal shares. Moving the worst-performing quartile of procurers to 75th percentile-effectiveness would reduce procurement expenditures by around 12 percent, or USD 10 billion each year—roughly 15 percent of the total amount the Russian state spends on health care. This would likely entail *better* spending quality too: the buyers in charge of procurement explain a quarter of the variation in spending quality, and price- and spending quality-effectiveness are positive correlated. Procurers' "type" thus appears to influence performance more than any multitasking incentives pulling price and spending quality apart.

We exploit our rich set of indicators on each procurer's auctions—measures of entry barriers chosen, how the auction was executed, procurer experience, etc—to select and explore the 30 most predictive correlates of estimated effectiveness (see also Lacetera et al., 2016). Consistent with our model, we find that effective procurers set lower reservation prices, and attract and admit more applicants to their auctions. While some other measures of bureaucrat behavior also predict effectiveness—for example, low-price procurers attract a somewhat wider variety of bidders—they generally do so to a lesser extent, and a wide range of auction characteristics, including measures of corruption, do not.

The second part of the paper focuses on the implications of heterogeneity in policy *implementer* effectiveness for policy *design*—often a more feasible path to better performance than attempting to directly increase bureaucrats' productivity. We focus on the particular example of industrial policy implemented through public procurement: bid preferences benefitting domestically manufactured goods.

In our model, introducing bid preferences makes participation less attractive to foreign bidders and more attractive to locals. When state effectiveness is high, so is baseline participation and so preferences induce a modest decrease in participation. However, when state effectiveness is low, baseline participation is low and so is the likelihood that a local bidder who enters has to face a more efficient, foreign, bidder. Bid preferences then have a large impact on the likelihood that a local bidder can win the contract, leading to a significant increase in participation. Additionally, foreign bidders shade their bids upward to offset the bid penalty. The overall impact on prices paid combines these participation and bidding responses with the mechanical effect of paying less to foreign winners. We show that the ultimate price effect depends negatively on baseline state effectiveness: effective buyers see performance worsen and vice versa.

We identify the impact of the bid preference regime using a generalized difference-in-differences approach that takes advantage of the fact that preferences apply to an evolving set of goods and are in effect for only parts of each year. Our results reveal that, *on average*, bid preferences achieve the Russian government's goal of channeling demand to domestic manufacturers, and do so at no cost to the government. If anything, average prices paid decrease slightly.[6]

[6]The contrast between this average treatment effect and that of similar policies in more effective

To test our model's heterogeneous treatment effect predictions, we interact the bid preference regime with our estimates of the effectiveness of the implementing bureaucrats. The small negative average effect on prices masks considerable heterogeneity. Our estimates imply savings of 12 vs. 0.7 percent when the policy is implemented by the least vs. the most effective quartile of bureaucrats, and that prices increase for some of the *most* effective bureaucrats (as has been shown for similar policies implemented in the U.S.).[7] We also find that characteristics of the procurers and how they initiate purchases that predict effectiveness in a constant policy regime—part one of our analysis—also predict how the policy affects prices paid when implemented by a given procurer. This suggests that policy changes can markedly affect state productivity even absent significant changes in policy implementer behavior.[8]

Overall, this paper demonstrates that state effectiveness is to a large extent embedded in the individuals and organizations of the bureaucratic apparatus, and that tailoring the design of policy to the implementing bureaucracy can partly offset the costs of bureaucratic ineffectiveness.

The first of two strands of literature on state effectiveness we contribute to focuses on individuals and the incentives they face as sources of productivity (see, among many others, Dal Bo, Finan and Rossi, 2013; Duflo et al., 2013, 2018; Bertrand et al., 2020; Khan, Khwaja and Olken, 2016, 2018; Rasul and Rogger, 2018).[9] We quantify, for the first time, the "macro" importance of the bureaucracy for public sector output—the share of overall variation in performance explained by bureaucrats *relative to (all) other contributors.* We sidestep concerns about multitasking and unobserved dimensions of performance by developing a new approach to measuring task-specific productivity and in parallel analyzing spending quality as a separate outcome.[10]

states (see e.g. Marion, 2007; Krasnokutskaya and Seim, 2011) suggests that industrial policies in public procurement may be more successful in countries with *low* average bureaucratic effectiveness. This foreshadows our next findings.

[7]In the pharmaceuticals sample, where we observe goods' origin, we also find that purchases administered by ineffective bureaucrats see a bigger increase in the probability that an auction is won by a supplier selling locally manufactured goods when bid preferences apply, consistent with our theoretical framework.

[8]We also find that additional features downstream in procurement processes—characteristics of the auction itself and the supplier—become important under bid preferences.

[9]Jones and Olken (2005); Xu (2018) study how public sector leaders and politicians matter for aggregate economic outcomes. In addition to Bandiera, Prat and Valletti (2009); Ferraz, Finan and Szerman (2015)—who, like us, focus on purchases of off-the-shelf goods—Lewis-Faupel et al. (2016); Coviello et al. (2017); Coviello, Guglielmo and Spagnolo (2018); Decarolis et al. (2018) also study state effectiveness in the context of public procurement. The innovative study by Decarolis et al. (2018) is especially related to this paper. The authors investigate how bureaucratic competence affects procurement outcomes in a setting where there are multiple dimensions to both competence and procurement outcomes, and find large effects.

[10]We thus avoid the limitations that arise from comparing workers and/or organizations (e.g. firms) (i) engaging in different or competing activities and/or (ii) based on wages and profits. The seminal work of Abowd, Kramarz and Margolis (1999); Abowd, Creecy and Kramarz (2002) spawned a large empirical literature using employer–employee matched datasets to address a range of important questions in labor economics (see, among many others, the papers cited in footnote 4, and also Bertrand and Schoar (2003) and the literature that followed on CEO effects). Wages do not necessarily reflect productivity (Card, Cardoso and Kline, 2016), but are important objects in and of themselves. Existing applications of

The second strand focuses on how public policy design should be tailored to context (see e.g. Laffont, 2005; Best et al., 2015; Duflo et al., 2018; Hansman, Hjort and León, 2019). The fact that *implementation* of policy is delegated to bureaucracies is often overlooked. Bureaucracies differ in effectiveness across contexts. We provide tools for measuring the performance of a bureaucracy and show that effectiveness affects the relative costs and benefits of different policies (see also Dehejia, Pop-Eleches and Samii, forthcoming).[11] We are not aware of prior studies that estimate treatment effects conditional on an unobserved characteristic such as effectiveness (see e.g. Heckman and Smith, 1997; Angrist, 2004, for discussion of the estimation of treatment effects conditional on observed characteristics).

I Public Procurement in Russia

A A decentralized system with centralized rules

Public procurement comprises 10 percent of Russia's non-resource GDP. In 1991, it created an extremely decentralized procurement system (see e.g. Enikolopov and Zhuravskaya, 2007). Each government entity has the legal authority to make its own purchases and there are no centralized purchases (such as framework contracts). Conversely, a federal law provides the legal framework for all procurement purchases above USD 35,000 for all levels of government (Yakovlev, Demidova and Balaeva, 2010).

We focus our analysis on electronic auctions—the most common vehicle, used for 53.5 percent of purchasing during our 2011–2016 data period—so as to study bureaucrats and organizations performing exactly the same task. Auctions are conducted through one of five independent web platforms. At the time of the auction, only the platform knows the identities of the bidders, making it possible to conduct auctions in which the bidding firms are anonymous to the procurers.

Appendix Figure H.1 traces out the steps involved together with the number of purchases that followed each path to contracting. The auction announcement is drawn up by a procurement officer. It contains technical details on the item(s) to be purchased (from clients), a maximum price, the required security deposit (between 0.5 and 5 percent of the maximum price), other participation requirements, and the auction date. Suppliers can then prepare a two-part application. The first part describes the good(s) that they are offering to fulfill the procurement

the AKM method have used samples that include workers performing many different tasks. Carneiro, Guimarães and Portugal (2012) show the potential importance of accounting for differences in tasks. On the organization/firm side, conventional methods estimate productivity from revenue or profits data and thus risk conflating productivity itself with mark-ups and quality differentiation (see e.g. Goldberg and De Loecker, 2014).

[11]Our findings resonate with those of the first studies to compare experimentally identified program effects across branches of companies or private-versus-public status of the implementing agency (see Bold et al., 2018; Allcott, 2015).

order. The second part—which cannot be accessed by the procurers until the auction is concluded—contains information on the supplier itself (name etc.).

A five-member commission, including the purchasing bureaucrat and organization, oversees the purchase. They receive and evaluate the anonymized first part of each application before the auction. The purchasing bureaucrat directs the commission's review to deny applications from suppliers that are not accredited, cannot pay the deposit, or whose proposal does not comply with the requested specifications.[12] If only one supplier is approved, the auction is declared "not held" and its contract is drawn up at the maximum price. This is relatively common, occurring in 1.4 million (22 percent of) cases. If there are no approved applicants, the purchase is cancelled (13 percent of purchases).

If more than one supplier is approved, the auction is held. Approved suppliers are assigned a number and remain anonymous. They log in to the platform and participate in a descending, open-outcry auction. Following the conclusion of the auction, the commission receives and reviews the second part of the applications. These contain the identifying information of the participants, but they cannot be linked to their bids. The commission checks the suppliers' accreditations, licenses, names, registration, and tax ID numbers. These are verifiable so there is little scope for subjective judgment at this stage. The contract is signed with the approved bidder who submitted the lowest bid.

B The role of bureaucrats and organizations in procurement

The labor market for Russian procurement officers resembles that for private sector jobs. Interested individuals seek out educational and employment opportunities in decentralized markets as in the private sector, creating labor market churn from procurement officers' and their employers' job search.[13] The Russian government does not educate bureaucrats, nor does it operate a centralized civil service administration to recruit, train, or assign public servants to postings (Barabashev and Straussman, 2007). In all cases we are aware of, procurement bureaucrats are paid a flat salary.

Purchases are made for the public entity that pays for and uses the goods; an *organization*. It may, for example, be a school, hospital or ministry, at the municipal, regional or federal level. To make a purchase, the organization must work with a procurement officer—individual *bureaucrats*. Together, the organization and bureaucrat (the *procurers*) acquire the good according to the centrally set rules, and at the lowest possible price. Any policy goals the central government may have, such as influencing which types of goods or firms win contracts, manifest themselves in the rules followed by all procurers. Conditional on following

[12] The platform accredits suppliers that are not in a state of bankruptcy; do not have substantial unpaid taxes; and are not listed in a registry of suppliers who have violated procurement rules during the last two years.

[13] Examples of private academies offering trainings on procurement include ArtAleks http://artaleks.ru/ and the Granit Center http://www.granit.ru/. The primary prerequisites are a legal education, management experience, and knowledge of current procurement laws.

those rules, procurers' only mandate is to pay the lowest possible price. For any given rules, the price paid is thus the primary measure of how effective procurers have been at implementing the government's procurement policy.

Bureaucrats can either be "in-house" (employees of the organization) or "external".[14] This means that we observe bureaucrats working with more than one organization (and vice versa) for two distinct reasons. The first is that bureaucrats change employers. The other is that external bureaucrats may conduct purchases with multiple organizations, and organizations may work with multiple external bureaucrats. On average, bureaucrats in our data are observed working with 5.2 organizations, and organizations with 4.8 bureaucrats. This high degree of churn is a powerful source of variation for this paper's empirical exercise.[15]

Since 2014, the division of labor between a procuring organization and an external procurement officer has been specified by law. The organization submits all technical documentation, and chooses and justifies the maximum price. The pair then together designate the commission to oversee the auction process. The bureaucrat manages consultations with specialists, collects information needed to design the tender, and works with the committee to conduct the first stage review, the auction, and the second stage review. The organization then signs the contract with the winner and verifies delivery. The same or a similar division of labor applies when in-house bureaucrats are used, and before 2014.

C Preferences for domestically manufactured goods

During our study period (2011–16), certain goods manufactured in Russia received a 15 percent bid preference for parts of each year. Where preferences are in place, if at least one bidder offers foreign-made goods and at least one offers locally manufactured goods, a bidder offering foreign-made goods only receives 85 percent of her final bid as the contract price.[16]

Each year from 2011 to 2014 a list of good categories for which a preference for domestic goods applied was drawn up.[17] The presidential order defining the

[14]Each regional authority sets rules dictating the type of bureaucrat used for each type of purchase, as defined by the maximum price and the nature of the item. External procurement agencies can be organized by a given authority (e.g. an education or health ministry), at the federal, regional, or municipal level. Part of the motivation for creating such agencies was to allow organizations purchasing similar goods to join forces and achieve lower prices. In practice, the decentralized management of procurement in Russia and coordination required means that joint purchases are very rare. Note that we control for the factors that authorities use to determine whether an in-house or external bureaucrat is used—the type of good and/or maximum allowable price of the contract—in our empirical analysis.

[15]Our setting features more turnover than would be observed in comparable private sector labor markets. German workers e.g. work at an average of 1.19 firms over the period 2002–2009 (authors' calculations based on Card, Heining and Kline, 2013).

[16]When the law is active, preferences formally apply to goods for which "the cost of goods produced in the territory of Russia, Belarus, and Armenia exceeds 50% of the total cost". Incorrect reporting of origin country may occur, but we found no coverage to suggest that such manipulation is common.

[17]Preferences were first given to domestic manufacturers in 2008 to stimulate the economy during the financial crisis. The list of goods covered was slightly changed in 2009, before expiring completely on December 31, 2010. The government then adopted an annual approach to determining which goods were covered beginning in 2011.

list was passed in May or June and remained in effect until the end of the year, after which the preference ceased to operate until a new list had been created and approved the following year (except in 2015 and 2016, when the 2014 list was extended through 2016). Preferenced goods spanned many categories, including automobiles, clocks, various food products, medical equipment, pharmaceuticals, textiles and furs (see Table H.1). Procurers were required to publicly inform potential suppliers that the preference applied.

Our analysis of the bureaucratic apparatus's role in procurement performance restricts attention to the policy regime without preferences. In Section V we analyze impacts of the preferences.

II Data and Measurement of Procurement Performance

Since 2011, a centralized procurement website (`http://zakupki.gov.ru/`) has provided public information about all purchases (EIS, 2022). We use data from this website on the universe of electronic auction requests, review protocols, auction protocols, and contracts from January 1, 2011 through December 31, 2016. The data cover 6.5 million auction announcements for the purchase of 21 million items. However, purchases of services and public works are idiosyncratic, and do not lend themselves to our approach to measuring performance, and so we remove them, resulting in a final sample of 16 million purchases of homogeneous goods. Table 1 describes the sample.

To use this data to evaluate procurement performance, we must overcome two challenges. First, we need very precise measures of the items being procured to use prices paid as our main measure of performance. Section II.A describes our text-based item measures. Second, prices are not the only outcome that matters in public procurement. Sub-section II.B describes the additional data we bring in to study bureaucrats' and organizations' impacts on spending quality. We round out this section by discussing how corruption affects our performance measures (Sub-section II.C) and the additional purchase process data we use to study the correlates of procurement performance in Sub-section II.D.

A Prices as performance

Our main measure of performance is the price paid, holding constant the precise nature of the item being procured. Holding constant the item being procured is crucial to avoid conflating differences in prices paid with differences in the precise variety of item being procured. A great deal of previous research in economics has faced this challenge, but typically achieves within-category homogeneity at the cost of losing generality.[18] To avoid doing so, we use the text of the final

[18] Broadly, three approaches have been taken: using hedonic regressions to estimate consumers' demand for and/or suppliers' costs of producing good attributes when rich attribute data is available (see e.g. Bandiera, Prat and Valletti, 2009); using product codes provided by e.g. customs agencies to partition goods (see e.g. Rauch, 1999); or restricting attention to products that are by nature especially homogeneous (Syverson, 2004).

contracts, in which the precise nature of the good purchased is laid out. We classify purchases into narrow product categories within which quality differences are likely to be negligible using text analysis methods (see also Hoberg and Phillips, 2016).

Our method proceeds in three steps. First, we transform the good descriptions in contracts into vectors of word tokens. Second, we use the universe of Russian customs declarations to train a classification algorithm to assign goods descriptions to a 10-digit Harmonized System product code, and apply it to the good descriptions in our procurement data. Third, for goods that are not reliably classified in the second step, either because the goods are non-traded, or because their description is insufficiently specific, we develop a clustering algorithm to group good descriptions into clusters of similar "width" to the categories from the second step. Details are in Online Appendix A.[19]

To complement this approach, we collect additional data on purchases of pharmaceuticals, an especially homogeneous category of goods (Bronnenberg et al., 2015). Russia's government regulates the pharmaceutical market, compelling suppliers of certain drugs to register in a List of Vital and Essential Medicinal Drugs (LVEMD) (MinZrav, 2016). This list includes information on each drug's active ingredient, i.e. international nonproprietary name (INN); the manufacturer's name and location; date of registration; and maximum price. Matching the LVEMD to our data, we can construct a barcode-level classification of pharmaceuticals.[20] The pharmaceuticals subsample is summarized in column (4) of Table 1.

B Spending quality

Sourcing inputs at low prices is the primary goal of public procurement,[21] but it is not the only outcome that matters. Contracts should not need to be unduly renegotiated or terminated, and goods should be delivered as specified, without delays. These outcomes reflect the quality of public spending and may conflict with the goal of achieving low prices, creating a multi-tasking problem for buyers. If this problem is severe, then we may misclassify bureaucrats and organizations as high-performing if they achieve low prices but this is offset by poor performance on spending quality.

To address this, we build direct measures of spending quality and use them as

[19]Online Appendix A also analyzes the sensitivity of our main findings to the choices made when developing our text analysis methodology. As Figure D.1 and Table E.4 show, the findings are robust.

[20]We use fuzzy string matching to combine the contract data on medicines with corresponding entries in LVEMD using each drug's international brand (trademark) name, active ingredient (INN), dosage, active units, concentration, volume, and units. We restrict the Pharmaceuticals Subsample to purchases of drugs we can match to LVEMD. Failure to match can arise if a medicine is not considered "essential" or because insufficient information is available in the procurement contract.

[21]Article 1 of Federal Law 94 (FZ-94), which transformed the public procurement system in 2005, declares the aim of procurement as the "effective, efficient use of budget funds". The law also introduced minimum price as the key criterion for selecting winners for most types of selection mechanisms (Yakovlev, Yakobson and Yudkevich, 2011).

an additional outcome in our analysis. We combine six proxies for the quality of the non-price outcomes of a procurement purchase: the number of contract renegotiations, the size of any cost over-run, the length of any delays, whether the end user complained about the execution of the contract, whether the contract was contested and canceled, and whether the product delivered was deemed to be low quality or banned for use in Russia because it didn't meet official standards. The first five of these measures come from the zakupki data, while the last one is sourced from a civil society organization—clearspending.ru—that scrutinizes the government's spending and publishes infractions they detect.

We focus on these six measures as they capture outcomes of a procurement purchase as opposed to inputs into the process leading up to the award of a contract. These are events that happen after the contract is signed that may not be captured in the contract price, but which alter the benefit to the government of the purchase. As a result, they should be thought of as outcomes which bureaucrats and organizations may affect. To summarize them in a single number, we take the six and create an index of spending quality y_i as the average of the six quality proxies after standardizing each one to have mean zero and standard deviation one: $y_i = \frac{1}{6}\sum_{k=1}^{6}(y_i^k - \bar{y}^k)/\sigma^k$ (Kling, Liebman and Katz, 2007).[22]

C Corruption

Both public procurement and Russia are associated with widespread corruption (Transparency International, 2016; Szakonyi, 2018). By its very nature, corruption is unobserved, and so we must take care to ensure that our measures of performance are not tainted by corruption. Corruption can lead to low quality goods being purchased at high prices. However, since our performance measure—the price paid conditional on the good—carefully controls for the precise nature of the good that is ultimately purchased, it captures both high prices and low item quality.[23] The reforms that introduced electronic procurement in Russia also imposed strict requirements on government customers whereby the final contract could only be ratified for the amount publicly disclosed on the auction platform. Corruption therefore is hiding in plain sight, 'on-the-books' in the mounds of procurement data. Savvy journalists have built numerous investigations into the misuse of government funds by analyzing the publicly available zakupki dataset.[24]

The quality-adjusted price paid is an attractive measure of performance in the potential presence of unobserved corruption for a number of reasons. First, governments mandate that procurers target exactly this—the price paid for goods

[22]We also use the first principal component of the five proxies and show that our results are similar. We prefer this index since it does not take a stand on placing higher weight on some proxies than on others (see Kling, Liebman and Katz, 2007).

[23]Note that it is important here that we use the item described in the final contract rather than the tender documents to capture leakage between what the tender documents specify and what is ultimately delivered.

[24]See for example Tom Bergin, and Stephen Grey: "Opaque Middlemen Exact High Price in Russia's Deals with the West." *Reuters*, December 19, 2014.

of specified quality. Second, quality-adjusted prices are the relevant metric when policy-makers decide which services can be offered given costs. Finally, both high prices stemming from a lack of effort or ability and high prices stemming from corruption represent transfers between taxpayers and bureaucrats and as such have similar welfare implications.

Of course, the underlying source of ineffectiveness may have welfare implications for higher-order efficiency or equity reasons.[25] However, the above arguments hold irrespective of whether quality-adjusted price differences are due to corruption or "intrinsic" ineffectiveness, and so in the model and empirical analysis below, we remain largely agnostic about their relative contributions. In Sub-section IV.E we provide some evidence that corruption is probably not the primary driver of variation in bureaucratic effectiveness in Russia (see also Bandiera, Prat and Valletti, 2009).

D *Process measurement*

In addition to measuring the performance outcomes described above, we also want to paint a detailed picture of the inputs bureaucrats and organizations provide in the procurement process. To do this, we exploit the richness of the zakupi procurement data, which contains details of the entire procurement process. This allows us to measure things such as the extent to which buyers rush at the end of the fiscal year (Liebman and Mahoney, 2017), the reservation prices buyers set, the number of items they bundle into purchases, the number of bidders who apply; are accepted; and bid in the auction, the competitiveness of the auction, the experience and types of products the buyers buy etc.

We supplement it with data from two sources. First, we use data from clearspending.ru on how well the process is run (these include whether identifier codes in the tender documents are correctly filled out, whether the names of the products in document headings are correct, whether sufficient time is provided to prepare a bid, whether the contract specifies the contractors correctly etc.) (ClearSpending, 2022). Second, we match firms in the procurement data to the Russian State Statistics Agency's firm databases (Rosstat, 2022*b*,*a*) and the Bureau Van Dijk's *Ruslana* database (Ruslana, 2022), which together cover the vast majority of firms that file financial information. This allows us to measure the types of firms that bid on, and that win, contracts from different buyers. Table F.1 summarizes the large number of variables we use on procurers' purchasing processes.

[25]Such consideration could for example arise if the source matters for whether ineffectiveness affects efficiency by changing which firms win government contracts, or if transfers to taxpayers and bureaucrats are valued differently for equity reasons. These possibilities present an important avenue for future research.

III A Model of Procurement with Heterogeneous State Effectiveness

In this section we present a stylized model of public procurement. We model state effectiveness as costs imposed on potential sellers wishing to participate in public procurement and show how variation in these costs leads to variation in output—the prices paid, motivating our empirical analysis in Section IV. We also show how the introduction of bid preferences differentially affects procurement by bureaucracies with different levels of state effectiveness, patterns we test for in Sub-section V.B.

A Performance heterogeneity in a constant policy environment

Consider a pair of a bureaucrat and an end-user organization—jointly, a bureaucracy—wishing to purchase an item from a supplier through a second-price descending auction. State effectiveness affects the prices the government is able to achieve in two ways. First, by directly increasing suppliers' contract fulfillment costs $\bar{\theta}/\theta_i$. $\bar{\theta}$ is a common cost component with three parts: $\log\left(\bar{\theta}\right) = \mathbf{X}'\boldsymbol{\beta} + \alpha_\theta + \psi_\theta$. $\mathbf{X}$ are observable attributes of the item and α_θ and ψ_θ are the costs of satisfying requirements stipulated by bureaucrats and organizations, respectively. These may include the date and place of delivery, the size of the order, and other requirements that directly affect the cost of fulfilling the contract. $\theta_i \geq 1$ is a firm-specific productivity term.

Second, bureaucrats and organizations indirectly affect prices by adding specifications α_c and ψ_c that affect the cost to firms of participating in the procurement process. These may include deposits required, the time granted to prepare bids, the clarity of the tender documents, bribes paid to enter the auction, and any other specifications affecting the cost of bidding, but not of fulfilling the contract.

In the first stage of the procurement process, two firms—one local and one foreign—observe the specifications $\{\mathbf{X}, \alpha_\theta, \alpha_c, \psi_\theta, \psi_c\}$ and decide whether to pay a participation cost c_i to learn their productivity θ_i and enter the auction.[26] The foreign firm $i = F$ and the local firm $i = L$ differ in both their expected productivity and their participation costs. Productivities θ_i are independent and Pareto distributed with Pareto parameters δ_F and δ_L. Foreign firms have higher expected productivities $(\delta_F < \delta_L)$[27] but face higher participation costs: $c_i = \frac{\bar{\theta}}{1+\delta_i} - \frac{\bar{\theta}}{1+\delta_L}\sqrt{1-\alpha_c-\psi_c}$.[28] In the second stage, if only one supplier chose to enter the auction, she is awarded the contract at price $\bar{\theta}$. If neither supplier entered, the bureaucracy finds an outside supplier and awards her the contract at

[26] We assume that firms do not know their productivity when they decide whether to enter the auction, as in Samuelson (1985). A more general approach would allow firms to have a signal of their productivity before deciding on entry as in Gentry and Li (2014). This significantly complicates the analysis, but the qualitative conclusions are the same. A sketch of such a model is available from the authors upon request.

[27] This fact is well established in the literature on international trade (see e.g. Bernard et al., 2007)

[28] This functional form makes the expressions for profits and prices tractable. However, the qualitative results only require the participation costs to be increasing in α_c and ψ_c.

a price of $\bar{\theta}$.[29] Finally, if both suppliers enter, a descending, open-outcry auction takes place, which we approximate with a second-price sealed-bid auction (see e.g. Milgrom, 2004).

The suppliers choose their entry and bidding strategies to maximize expected profits. We outline the equilibrium here, relegating a detailed characterization and the proofs of propositions to Online Appendix B. Working backwards from the second stage, when both firms enter, it is a dominant strategy for bidders to bid their fulfillment cost since bidder valuations are independent (see e.g. Milgrom, 2004). The winner is the bidder with the lowest fulfillment cost and receives the contract at the other bidder's fulfillment cost. The participation decision depends on the size of the participation costs c_i. When participation costs are sufficiently small, both firms enter and the auction always takes place. For larger participation costs the equilibrium involves mixed strategies with entry probabilities q_i. We can summarize the equilibrium in the following proposition:

PROPOSITION 1: *In the Nash equilibrium of the auction, the bidders, $i \in \{F, L\}$ enter with probabilities $q_i = \sqrt{\kappa\left(1-\alpha_c-\psi_c\right)}$, where* $\kappa = \min\left\{\left[\left(1+\delta_F+\delta_L\right)/\left(1+\delta_L\right)\right]^2, 1/\left(1-\alpha_c-\psi_c\right)\right\}$. *Expected log prices are*

$$\mathbb{E}\left[\log\left(p\right)\right] = \log\left(\bar{\theta}\right) - \tfrac{q_F q_L}{\delta_F+\delta_L} = \mathbf{X}'\boldsymbol{\beta} - \tfrac{\kappa}{\delta_F+\delta_L} + \tilde{\alpha} + \tilde{\psi}, \tag{1}$$

where $\tilde{\alpha} = \alpha_\theta + \frac{\kappa}{\delta_F+\delta_L}\alpha_c$, and $\tilde{\psi} = \psi_\theta + \frac{\kappa}{\delta_F+\delta_L}\psi_c$. In equilibrium

1) *Bureaucracies that impose higher contract fulfillment costs α_θ, ψ_θ pay higher prices for otherwise identical goods.*

2) *Bureaucracies that impose higher participation costs α_c, ψ_c pay higher prices for otherwise identical goods, and also attract fewer bidders to auctions they run.*

Equation (1) shows how prices vary with with the costs imposed by bureaucrats ($\tilde{\alpha}$) and organizations ($\tilde{\psi}$) managing the procurement process, and forms the basis of our empirical approach.

B Policy change with heterogeneous state effectiveness: bid preferences

We now study the impact of introducing bid preferences favoring the locally producing bidder L. If the lowest-bid, winner of the auction is foreign, the contract price will only be $p = \gamma b_L$, where $\gamma < 1$, while a local winner receives the

[29] A more realistic assumption might be that auctions in which no firms enter have to be re-run at some cost. Our assumption makes the model static, simplifying the exposition. The qualitative results are unlikely to depend on this choice since no firms entering is more likely for low-effectiveness buyers (since, as discussed below, firms weigh entry costs against expected profits from participation and low-effectiveness buyers impose higher entry costs), and so this channel only adds to the additional costs that low-effectiveness buyers create.

undiscounted $p = b_F$. Otherwise the auction protocol is unchanged. Preferences make it optimal for bidder F to shade so that her contract price should she win is equal to her fulfillment cost $b_F = \bar{\theta}/\gamma\theta_F$. However, when her shaded bid would have no chance of winning ($\theta_F < 1/\gamma$), she drops out and the contract is awarded to bidder L.

The effects on prices depend on the balance of four effects. First, the penalty mechanically lowers prices in auctions with foreign winners. Second, local bidders, who are less productive on average, are advantaged in the auction, raising prices.[30] Third, since foreign bidders are less likely to win auctions, they are less likely to participate. Fourth, local bidders are emboldened to enter by their higher chance of winning the contract. The interesting cases arise when the preferences are strong enough that the effect on L's entry decision is considerable, but not so large as to make it very unlikely F can win the auction. Formally, we focus on the case when $\gamma^{-\delta_F} > 1 - \log\left(\gamma^{\delta_L}\right)$.[31] In this case, introducing bid preferences has heterogeneous effects depending on the effectiveness of the bureaucracy that we summarize in the following proposition:

PROPOSITION 2: *When $\gamma^{-\delta_F} > 1 - \log\left(\gamma^{\delta_L}\right)$, the introduction of bid preferences has different effects on three groups of bureaucracies differing in their effectiveness.*

1) *For bureaucracies with $\alpha_c + \psi_c \leq \underline{c}$, prices rise, the expected number of bidders is unchanged, and the probability that bidder L wins the contract at auction increases;*

2) *For bureaucracies with $\underline{c} < \alpha_c + \psi_c \leq \bar{c}$, prices rise, the expected number of bidders falls, and the probability that bidder L wins the contract at auction decreases;*

3) *For bureaucracies with $\bar{c} < \alpha_c + \psi_c$, prices fall, the expected number of bidders increases, and the probability that bidder L wins the contract at auction increases. The probability that bidder L wins the contract at auction increases by more than in case 1.*

The thresholds $\underline{c}$ and $\bar{c}$ are defined by

$$\underline{c} = 1 - \left(\tfrac{1+\delta_L}{1+\delta_F}\left(1-\gamma^{1+\delta_F}\right) + \tfrac{1+\delta_L}{1+\delta_F+\delta_L}\gamma^{1+\delta_F}\right)^2 \qquad \bar{c} = 1 - \left(\tfrac{1+\delta_L}{1+\delta_F+\delta_L}\gamma^{\delta_F}\right)^2.$$

For effective bureaucracies that impose low participation costs on potential bidders ($\alpha_c + \psi_c \leq \underline{c}$), preferences do not deter foreign firms from entering the

[30] There is extensive evidence that exporters are more productive than other firms, see e.g. Bernard et al. (2007).

[31] Essentially, this condition requires that δ_L not be too much larger than δ_F. If this is violated, even with the preferences, the probability the local bidder wins is still very small and so when there is an auction the foreign bidder still wins but has her bid penalized lowering final prices even when the bureaucracy is very effective.

auction, but the local bidder is more likely to win, and the less aggressive bidding by the foreign bidder raises expected prices. For bureaucracies with intermediate effectiveness ($\underline{c} < \alpha_c + \psi_c \leq \bar{c}$), foreign bidders no longer find it profitable to enter. Since only the local bidder enters, the auction does not take place and the local firm gets the contract at the maximum price $\bar{\theta}$. Finally, when bureaucracies impose high participation costs ($\bar{c} < \alpha_c+\psi_c$), the increase in bidder L's willingness to enter is larger than the decrease in bidder F's willingness to enter, increasing the probability of both bidders entering and the auction taking place, lowering expected prices. Moreover, the entry effect is larger than the increase in prices caused by the changes in the bidding behavior in the auction, resulting in an overall decrease in expected prices.

Proposition 2 makes three predictions about heterogeneity in the impact of bid preferences. First, bureaucracies that pay higher prices when there are no bid preferences—which Proposition 1 shows is associated with higher participation costs—should experience price *decreases*, while bureaucracies that pay lower prices absent the bid preferences experience price *increases*. Second, the average number of participants in procurement processes should increase for bureaucracies that pay higher prices when there are no bid preferences. Third, we should see that the probability that an auction is won by a bidder offering to supply locally manufactured goods increases by more for bureaucracies that pay higher prices when there are no bid preferences. These are the patterns we test for in Sub-section V.B

IV How Important is a Good Bureaucracy?

In this section we estimate the extent to which procurement effectiveness can be attributed to the individuals and organizations in the bureaucracy. We extend the method pioneered by Abowd, Kramarz and Margolis (1999) exploiting *switchers*—bureaucrats who make purchases with multiple organizations, and organizations who make purchases with multiple bureaucrats—for identification.

A Identifying the effectiveness of individuals and organizations

We start by showing that bureaucrat-organization switches identify the causal impact of the individual in charge and the organization he or she works with on the purchase price. We use an event study analysis to study the dynamics of prices paid by organizations around the time that they switch the bureaucrat they work with. This happens frequently in Russia. As detailed in Table D.1, we observe 65,000 events in which organizations switch bureaucrats, with an average of 45 observations per event.

We define an event as chronological pairs of employment spells involving the same organization but two different bureaucrats. Figure 1 shows how prices change around such events. Each of the two employment spells is a sequence of at least two weeks less than 400 days apart in which a bureaucrat-organization pair makes purchases together. The horizontal axis displays event time, i.e. purchase weeks. The vertical axis displays the average quality-adjusted prices paid in a given week. The figure shows the evolution of prices paid by buyers starting with a bureaucrat in the top or bottom quartile of effectiveness, which we define using purchases made by the bureaucrats involved in the event, but that are not included in the event itself. Specifically, we use the average quality-adjusted price they pay in purchases made for *other* organizations they work with during the half-year that the spell ends (for the earlier spell) or starts (for the later one), akin to Card, Heining and Kline (2013).[32]

Four key findings emerge from Figure 1. First, quality-adjusted prices paid change sharply, and in the expected direction, precisely when an organization switches to a less or more effective bureaucrat. The estimates suggest that an organization switching from a worst quartile-bureaucrat to a best quartile-bureaucrat on average experiences an 18 percent decrease in prices paid. Second, the figure shows no sign that performance is improving in organizations that subsequently switch to a better bureaucrat, and vice versa.[33] This suggests that drift in effectiveness and switches are uncorrelated. Third, we do not see a systematic dip or spike in performance before a bureaucrat switch, indicating that switches are not driven by temporary improvements or deteriorations in performance. Fourth, the price changes associated with switching bureaucrats appear symmetric: organizations switching from a bureaucrat in the best quartile of average prices to one in the worst quartile experience a price *increase* of similar magnitude to those switching in the other direction. In Online Appendix D we show that these patterns are robust to changing a series of choices made in constructing the event studies.

Taken together, the evidence in this sub-section suggests that the thousands of quasi-experiments arising from organizations switching bureaucrats and vice versa in Russian public procurement can be used to estimate specific procurers' causal impact on performance, and that this impact is large.[34]

B Variance decomposition method

We now aggregate the causal effects of specific bureaucrats and organizations from Sub-section IV.A into estimates of the share of sample-wide variation in procurement performance bureaucrats and organizations as a whole explain. To

[32]We quality-adjust prices by regressing them on log quantity, good fixed effects, month fixed effects, interactions between 2-digit HS product categories, years, regions, and lot size (as detailed in the next sub-section). Table D.1 highlights that the number of switches used to construct each quartile-to-quartile plot in Figure 1, and the average number of purchases observed for each bureaucrat-organization involved, are symmetric both around the events, and across quartile-to-quartile plots. The table also displays the average number of calendar weeks between each purchase week on the x-axis of Figure 1.

[33]More formally, of the sixteen groups formed by the possible trajectories between the four quartiles of bureaucrat effectiveness, we are unable to reject the null of no pre-trend in ten. Of the remaining six, five have pre-trends that point in the opposite direction of this concern.

[34]We also construct analogous event study figures for organizations and bureaucrats switching from purchasing one type of *good* to another. The results are in Figure D.4. Each event study shows the same general patterns as in Figure 1.

do so we first extend the method pioneered by Abowd, Kramarz and Margolis (1999) to study wage dispersion in the private sector, and then show how to correct for sampling error to form predictions of the impact of each bureaucrat and organization on prices paid. We use these predictions to examine the mechanisms through which procurers affect prices in Sub-section IV.E and how bureaucratic effectiveness impacts the way policies map into public sector output in Section V.

We model the price paid for item i procured by organization j and bureaucrat $b(i, j)$ as a function of item attributes $\mathbf{X}_i$, a price premium due to the bureaucrat $\tilde{\alpha}_{b(i,j)}$, and a price premium due to the organization $\tilde{\psi}_j$. As the theoretical framework in Section III shows, these price premia can be thought of as a reduced form for the impact on prices of the participation costs that bureaucrats and organizations of different levels of effectiveness impose on suppliers. The log unit price paid for an item is

$$p_i = \mathbf{X}_i \boldsymbol{\beta} + \tilde{\alpha}_{b(i,j)} + \tilde{\psi}_j + \varepsilon_i \tag{2}$$

To control flexibly for the item being purchased, $\mathbf{X}_i$ includes log quantity, good and month fixed effects, and interactions of 2-digit HS product categories, years, regions, and lot size.[35]

Identifying the bureaucrat and organization premia is made possible by the switches we documented in Sub-section IV.A. As Abowd, Creecy and Kramarz (2002) show, individual and organization effects are only identified *within* sets of organizations connected by individuals moving between them.[36] However, such switches do not connect all bureaucrats and organizations that conduct procurement in Russia. Our data contain 616 connected sets. This relatively large number comes about for several reasons. First, focusing on bureaucrats performing a *single task*, rather than comparing many types of workers through their wages—the approach taken in existing related work—limits connectedness. Second, workers change employers less often in the public than in the private sector. Finally, the decentralized nature of Russian procurement means that some geographically remote organizations do not have bureaucrat links to other organizations.

To form our Analysis Sample, we focus on connected sets containing at least

[35] By lot size we mean the maximum allowable price for all items to be purchased in the auction. We divide this price into bins to allow our estimates of effectiveness to capture the impact on prices of the procurers' choice of the exact maximum price posted. The interactions help address e.g. concerns that systematic spatial variation in the average prices of different types of goods—Russian regions are highly heterogeneous (Enikolopov and Zhuravskaya, 2007; Yakovlev and Zhuravskaya, 2014)—in combination with differences across procurers in items purchased, confound our estimates of effectiveness. Hereafter we refer to the goods categories constructed using the method described in Sub-section II.A as "goods".

[36] More precisely, within each connected set s containing $N_{b,s}$ bureaucrats and $N_{j,s}$ organizations, we can identify at most $N_{b,s} + N_{j,s} - 1$ linear combinations of bureaucrat and organization fixed effects. In fact, we estimate models with three sets of high-dimensional fixed effects, for bureaucrats, organizations, and goods (the models also contain month dummies to control for common time trends, but there are few enough of these month effects such that "month-connectedness" is not an issue). To our knowledge, identification results for models with more than two sets of fixed effects are not yet available (Gaure, 2013), however our focus is on the estimates of only two of the three dimensions–the bureaucrat and the organization effects.

three bureaucrats and organizations after we make the following restrictions. We remove any procurer pairs that only ever occur together (as in this case it is not possible to distinguish bureaucrat and organization effects), and similarly for bureaucrat-good pairs and organization-good pairs as well as any levels of our control fixed effects that only appear once in the data. We also require that all bureaucrats and organizations in the Analysis Sample make at least five purchases. Table 1 compares the full sample and the Analysis Sample. The organizations in the Analysis Sample are less likely to be federal, but their purchases are of similar size and quantity to those in the full sample.[37] Overall the sample we use for analysis appears to be fairly representative.[38]

To proceed, we normalize the $\tilde{\alpha}_{b(i,j)}$ and $\tilde{\psi}_j$ to have mean zero in each connected set and augment (2) to include intercepts $\gamma_{s(b,j)}$ for each connected set:

$$p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i \tag{3}$$

In Online Appendix C, we show that while the $\tilde{\alpha}$s and $\tilde{\psi}$s in equation (2) are not identified, the αs, ψs and γs in equation (3) are. These are related to the underlying bureaucrat and organization effects as follows: $\alpha_b = \tilde{\alpha}_b - \overline{\alpha}_{s(b)}$, $\psi_j = \tilde{\psi}_j - \overline{\psi}_{s(j)}$, and $\gamma_{s(b,j)} = \overline{\alpha}_{s(b,j)} + \overline{\psi}_{s(b,j)}$, where $\overline{\alpha}_{s(b)}$ is the mean bureaucrat effect in the connected set containing bureaucrat b, and similarly $\overline{\psi}_{s(j)}$ is the mean organization effect in organization j's connected set.[39]

We can use equation (3) to decompose the variance of prices into its constituent parts using

$$\begin{aligned}\mathsf{Var}(p_i) = {} & \mathsf{Var}\left(\alpha_{b(i,j)}\right) + \mathsf{Var}\left(\psi_j\right) + 2\mathsf{Cov}\left(\alpha_{b(i,j)}, \psi_j\right) \\ & + 2\mathsf{Cov}\left(\alpha_{b(i,j)} + \psi_j, \gamma_{s(b,j)} + \mathbf{X}_i\boldsymbol{\beta}\right) + \mathsf{Var}\left(\gamma_{s(b,j)} + \mathbf{X}_i\boldsymbol{\beta}\right) + \mathsf{Var}\left(\varepsilon_i\right)\end{aligned} \tag{4}$$

all of which can be identified. Since $\mathsf{Var}\left(\alpha_{b(i,j)}\right)$ and $\mathsf{Var}\left(\psi_j\right)$ are variances within connected sets, they are lower bounds on the underlying variances of bureaucrat and organization effects.[40]

We can obtain unbiased estimates of procurer effects using OLS under the assumption that the residuals in (3) are uncorrelated with the identity of the bureaucrat or organization making a purchase (conditional on $\mathbf{X}_i$). There are two principal reasons this might not be the case. First, it could be that prices

[37] We find that bureaucrats at federal agencies switch jobs less often, since there is more scope for both horizontal and vertical mobility within these larger organizations.

[38] In Table E.3 we show that our results are robust to using only the largest set of connected organizations. Table E.2 compares the Analysis Sample to its largest connected set.

[39] Faced with this issue, previous work on firms and workers has tended to restrict attention to the largest connected set, normalizing an arbitrary firm effect to 0, and estimating unconditional variances. An exception is Card, Cardoso and Kline (2016).

[40] Formally, $\mathsf{Var}\left(\tilde{\alpha}_b\right) \equiv \mathbb{E}\left[\mathsf{Var}\left(\tilde{\alpha}_b | s(b)\right)\right] + \mathsf{Var}\left(\mathbb{E}\left[\tilde{\alpha}_b | s(b)\right]\right) = \mathsf{Var}\left(\alpha_b\right) + \mathsf{Var}\left(\mathbb{E}\left[\tilde{\alpha}_b | s(b)\right]\right) \geq \mathsf{Var}\left(\alpha_b\right)$. Similarly, $\mathsf{Var}\left(\tilde{\psi}_j\right) = \mathsf{Var}\left(\psi_j\right) + \mathsf{Var}\left(\mathbb{E}\left[\tilde{\psi}_j | s(j)\right]\right) \geq \mathsf{Var}\left(\psi_j\right)$.

change around the time bureaucrats move across organizations or vice versa, for reasons unrelated to the switch. However, as shown in Sub-section IV.A, we do not see evidence of such pre-trends.

Second, equation (3) assumes that prices are log-linear in the procurer effects—an assumption about the degree of complementarity between the bureaucrat and the organization working on a purchase and associated sorting patterns. If the model is misspecified, then the omitted complementary terms are a component of the residuals in (3).[41] These complementarities may be correlated with the identity of the bureaucrat or organization making a purchase if, for example, organizations recruit bureaucrats who specialize in particular goods. Then estimates from (3) would recover a mixture of the true effect and the average complementarity of bureaucrat-organization matches.

Such sorting would imply that organizations switching from bureaucrats who pay high prices to bureaucrats who pay low prices enjoy larger decreases than the price increase suffered from moving in the opposite direction. Organizations hiring a low-price bureaucrat benefit from *both* a lower average price and an improved match effect, and organizations hiring a high-price bureaucrat lose from the lower average price but benefit from an offsetting improved match effect. We see no evidence of such patterns in Figure 1.[42] The symmetry of the event study evidence indicates that omitted complementarities are unlikely to bias our estimates. Online Appendix E.E1 provides further tests for misspecification.

We use a large sample of public procurers, but nevertheless, our estimates need not be consistently estimated, even if they are unbiased. Consistency of the estimated fixed effects requires that the number of observations *on each group* tends to infinity (Lancaster, 2000). Our data contains 284,710 bureaucrat-organization pairs and an average of 40 observations per pair, so we cannot be confident a priori that the error in the bureaucrat and organization effect estimates has asymptoted to zero, particularly for the less frequently observed pairs. Moreover, since we are estimating two sets of fixed effects, the problem is compounded if the network features too few switches. Such *limited mobility bias* results in a spurious negative correlation between the two dimensions of estimated fixed effects (Andrews et al., 2008). Each connected set in our data is densely connected—we observe bureaucrats working with 5.2 organizations on average, and organizations with 4.8 bureaucrats—but limited mobility bias may still be a concern.[43]

We address these sampling error issues in three ways. First, we bootstrap to estimate standard errors for our variance decomposition.[44] Second, we take a

[41]Our identifying assumption does not rule out effective bureaucrats and organizations matching with each other.

[42]If anything, the price decreases when organizations switch to lower average-price bureaucrats in Figure 1 appear slightly *smaller* than the corresponding increases when organizations switch to higher average-price bureaucrats.

[43]Moreover, in 76% of organizations, *all* the bureaucrats they work with are switchers (work with multiple organizations). Similarly, for 94% of bureaucrats, all the organizations they work with are switchers. This is reassuring since it is these switches that allow us to identify their effects.

[44]We construct partial residuals $\epsilon_i = p_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$ and randomly resample the residuals, stratifying by

non-parametric, split-sample approach to estimating the variance components in (4), akin to Finkelstein, Gentzkow and Williams (2016) and Silver (2016). We randomly split our sample in half, stratifying by bureaucrat-organization pair. We then estimate equation (3) separately on each sample, yielding two estimates ($k = 1, 2$) for each bureaucrat ($\hat{\alpha}_b^k$), organization ($\hat{\psi}_j^k$), and connected set ($\hat{\gamma}_s^k$) effect. Both estimates are estimated with error, but the errors in the two estimates should be uncorrelated, so we can create split-sample estimates of the variance decomposition terms as follows: $\widehat{\mathsf{Var}}^{SS}(\alpha_b) = \mathsf{Cov}\left(\hat{\alpha}_b^1, \hat{\alpha}_b^2\right)$, $\widehat{\mathsf{Var}}^{SS}(\psi_j) = \mathsf{Cov}\left(\hat{\psi}_j^1, \hat{\psi}_j^2\right)$, $\widehat{\mathsf{Var}}^{SS}(\gamma_s) = \mathsf{Cov}\left(\hat{\gamma}_s^1, \hat{\gamma}_s^2\right)$, and $\widehat{\mathsf{Var}}^{SS}(\alpha_b + \psi_j) = \mathsf{Cov}\left(\hat{\alpha}_b^1 + \hat{\psi}_j^1, \hat{\alpha}_b^2 + \hat{\psi}_j^2\right)$.

Third, we adopt two shrinkage approaches to create predictions of each bureaucrat and each organization effect. The variance in our estimated fixed effects comes from two sources: the true, signal variance in bureaucrats' and organizations' effects, σ_α^2 and σ_ψ^2 respectively, and sampling error with variances σ_μ^2 and σ_ω^2. Bootstrapping the estimation of equation (3) yields estimates of the variance of the sampling error which we use to perform a standard shrinkage procedure for the bureaucrat and organization estimates separately, as is common in studies of teacher value-added (see e.g. Kane and Staiger, 2008; Chetty, Friedman and Rockoff, 2014*a*).[45] To address limited mobility bias, we extend the shrinkage approach used in existing work to explicitly account for the correlation between the estimation errors of the bureaucrat and organization effects. Our bootstrap also provides estimates of the covariance of all the estimation errors which we use to form minimum mean-squared error predictions of the full vector of bureaucrat and organization effects.[46] We label this method "covariance shrinkage". It yields our preferred estimates of the price variance decomposition in equation (4).[47]

bureaucrat-organization pair to preserve the match structure of the observations. We then re-estimate the bureaucrat and organization effects. We repeat this procedure 100 times, and use the distribution of the estimates to compute standard errors. This procedure does not fully account for uncertainty arising from the data's match structure and finite sample correlations between bureaucrat and organization assignment and $\mathbf{X}$, but is computationally feasible.

[45]Formally, we find $\lambda_b = \arg\min_{\tilde{\lambda}} \mathbb{E}\left[\alpha_b - \tilde{\lambda}\hat{\alpha}_b\right] = \sigma_\alpha^2 / \left(\sigma_\alpha^2 + \sigma_{\mu_b}^2\right)$, and analogously for λ_j. Our shrinkage estimators replace these terms with their sample analogues $\hat{\alpha}_b^{Sh} = \hat{\lambda}_b\hat{\alpha}_b$ and $\hat{\psi}_j^{Sh} = \hat{\lambda}_j\hat{\psi}_j$.

[46]Formally, we seek the linear combination of the full vector of fixed effects that minimizes the expected mean-squared error of the predictions. Denoting the vector of estimated bureaucrat and organization fixed effects by $\hat{\boldsymbol{\theta}}$ and the matrix of weights by $\boldsymbol{\Lambda}$, the objective is $\min_{\boldsymbol{\Lambda}} \mathbb{E}\left[\left(\boldsymbol{\theta} - \boldsymbol{\Lambda}\hat{\boldsymbol{\theta}}\right)'\left(\boldsymbol{\theta} - \boldsymbol{\Lambda}\hat{\boldsymbol{\theta}}\right)\right]$, which has solution $\boldsymbol{\Lambda}^* = \mathbb{E}\left[\boldsymbol{\theta}\hat{\boldsymbol{\theta}}'\right]\left(\mathbb{E}\left[\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}'\right]\right)^{-1}$. Replacing the expectations with their sample analogues yields the shrinkage matrix $\hat{\boldsymbol{\Lambda}}^* = diag\left(\hat{\sigma}_\alpha^2, \hat{\sigma}_\psi^2\right)\left(diag\left(\hat{\sigma}_\alpha^2, \hat{\sigma}_\psi^2\right) + \boldsymbol{\Sigma}\right)^{-1}$, where $\boldsymbol{\Sigma}$ is the covariance matrix of the bootstrap estimates and $diag\left(\hat{\sigma}_\alpha^2, \hat{\sigma}_\psi^2\right)$ is the diagonal matrix with $\hat{\sigma}_\alpha^2$ in entries corresponding to entries for bureaucrats in $\boldsymbol{\theta}$ and $\hat{\sigma}_\psi^2$ in entries corresponding to organizations.

[47]We thus use "covariance shrunk" estimates in our analysis of the determinants of bureaucratic capacity in Sub-section IV.E and the analysis of the effects of procurement policy changes in Section V. For computational reasons, we perform covariance shrinking separately in each connected set. Since the estimated fixed effects are all normalized to be mean zero within each connected set and by definition the observations are unrelated across connected sets, this is without loss.

C Results

Table 2 shows results from our variance decomposition (4). The first column shows estimates of the standard deviations using the raw fixed effects estimates from equation (3), while estimates from the split-sample approach are in Column (3). The corresponding standard errors are in columns (2) and (4). The results from the shrinkage and covariance shrinkage methods are in columns (5) and (6).

Three key findings emerge. First, bureaucrats and organizations are each important determinants of policy performance. After controlling for the good being purchased and the month of the purchase, the standard deviation of log unit prices is 1.280. Compared to this, the bureaucrat fixed effects have a standard deviation of 0.788 and the organization fixed effects' standard deviation is 0.927. The split-sample estimates in Column (3) are similar. The shrinkage methods in columns (5) and (6) deliver slightly smaller estimates of the bureaucrat and organization variances, but even the covariance shrinkage estimates imply large effects of bureaucrats and organizations on policy performance.

Second, the covariance shrinkage method shown in Column (6) appears to best deal with the finite-sample inconsistency of our estimates. The fixed effects, split-sample, and shrunk estimates all yield a negative estimate of the correlation between bureaucrat and organization effects.[48] However, our covariance shrinkage approach yields a more plausible estimate of the correlation of 0.311.[49] As a result, the covariance shrunk estimates of share of the variation in performance explained by bureaucrats and organizations—21 and 26 percent percent respectively—are our preferred estimates.

Third, the combined importance of bureaucrats and organizations for policy performance is large. Our estimates of the within-connected-set standard deviation of the combined bureaucrat and organization effects are consistent across the four methods, ranging from 0.66 for the split-sample approach down to our preferred estimate of 0.49 for the covariance-shrunk estimates—38 percent of the standard deviation of log unit prices. Overall, our estimates imply that bureaucrats and organizations jointly explain a remarkably large share of the variation in procurement effectiveness in Russia, of which about half in turn is due to bureaucrats and half to organizations.

The large estimates in Table 2 have correspondingly dramatic implications for the scope of potential savings from improving the effectiveness of the bureaucracy. One way to illustrate the magnitude is to consider simple counterfactual bureaucracies in which bureaucrats and/or organizations with low effectiveness are improved, for example through changes in recruiting, training of existing

[48] The same is found in many studies applying the AKM method to private sector wages. This led Andrews et al. (2008) to show that the AKM-estimated covariance term is downward biased (see Subsection IV.B) and to suggest a parametric correction. However, this parametric correction relies on homoskedasticity of the residuals, an unappealing requirement in our setting (see also Card, Heining and Kline (2013)).

[49] Recall that such assortative matching does not violate the no-sorting-on-match-effects assumption discussed above.

bureaucrats, or improved organizational management. Our estimates indicate that increasing the effectiveness of the lowest quartile of bureaucrats to the 75th percentile would save the Russian government 4.6 percent of annual procurement expenses. Moving all bureaucrats *and* organizations below 25th percentile-effectiveness to 75th percentile-effectiveness would save the government 13.9 percent of procurement expenditures. Annual procurement expenses are USD 86 billion, so this implies savings of USD 10 billion each year, or 0.7 percent of non-resource GDP (see Table H.2)—roughly one fifth, for example, of the total amount spent on health care in 2013 and 2014.[50]

Another way to illustrate the magnitude of our estimates is to compare them to existing estimates of the extent to which individuals and organizations affect output in other settings. Several studies are indirectly comparable. Studying front-line service providers in rich countries, Chetty, Friedman and Rockoff (2014*b*) find that increasing the performance of 5th percentile American grade 3–8 teachers to 50th percentile would increase the present value of their students' lifetime incomes by 2.76 percent, and Silver (2016) finds that improving the performance of American emergency room doctors by one standard deviation would decrease time-of-care by 11 percent. We find that the same (relative) improvement in performance among Russian procurement officers would lower prices paid by 33.0 and 30.2 percent respectively. In studies of workers in the private sector performing a simpler task, Mas and Moretti (2009) and Lacetera et al. (2016) find, respectively, that increasing performance by one standard deviation would decrease cashier processing times in a U.S. supermarket chain and increase the probability of cars being sold in U.S. used-car auctions by 11 and 4.3 percent, while in our case the improvement is 36.5 percent. [51]

D Robustness

We interpret the results in the previous sub-section as capturing the total, causal contribution of bureaucrats and organizations to the Russian state's effectiveness in procuring off-the-shelf goods. But are we adequately controlling for the precise item being purchased? And while prices paid are the primary metric of procurement effectiveness (see Sub-section II.B), they are not the only one—what about spending quality?

[50]Appendix Figure E.3 shows how these counterfactuals affect the distributions of effectiveness.

[51]Of course, (i) teachers and doctors may differ from procurement officers in the complexity of the job performed, motivations, and many other dimensions, while (ii) output is less easily measured and monitored in the public sector than among private sector cashiers and auctioneers so we expect greater scope for differences between bureaucrats. We are not aware of comparable estimates of the causal effects of workers and organizations on output in a low or middle-income country government context. We perform these calculations separately in each connected set and report the average, weighting by the number of items.

Like-for-like comparisons

If our goods classification based on contract texts is inaccurate, our estimates will conflate the true effects on prices with differences across bureaucrats and organizations in products bought. To probe this concern, we first show that our findings are similar in a sub-sample of goods that is by nature homogeneous—pharmaceuticals (see also Syverson, 2004; Bronnenberg et al., 2015). We create barcode-level bins for pharmaceuticals as described in Sub-section II.A and make the same connectivity restrictions as in the full sample to create an analysis sample. Columns (4) and (5) of Table 1 summarize the sample. Table 3 presents the results of re-estimating (3) on the pharmaceuticals sample. Naturally, since the sample is more homogeneous and our barcode product categories are very precise, the share of the variation in prices explained by the good fixed effects is larger than in the broader sample. However, of the remaining variation in policy performance, all but the covariance shrinkage method attribute 30–40 percent to the combination of bureaucrats and organizations.[52] This is strikingly similar to the 40 percent found in the broader analysis sample. This is also what our theoretical framework suggests we should see, since we model the fulfillment costs imposed by bureaucrats and organizations on suppliers as proportional costs.

Second, our results are robust to focusing on more homogeneous subsets of goods in our full sample. In Figure 2 we split the sample into quintiles of good homogeneity as defined by the commonly-used measure of scope for quality differentiation developed by Sutton (1998).[53] We then reestimate (3) on successive subsamples. As we move from right to left, we restrict the sample to more and more homogeneous goods. As expected, the overall variance of average prices paid, shown by the grey shaded areas, decreases with good homogeneity. However, as shown by the blue line, the estimated share of the variance explained by bureaucrats and organizations remains very similar across the columns. In Appendix Figure E.2 we repeat this exercise using an alternative measure of scope for quality differentiation developed by Khandelwal (2010) and find the same result.[54]

Third, the right-most bar in Figure 2 shows that the results from our variance decomposition exercise are also essentially unaffected if we restrict the sample to items the text-based classification method is confidently able to assign a 10-digit Harmonized-System product code to.[55]

[52]The covariance shrinkage method is less reliable in this sample since we have an order of magnitude fewer observations per connected set (an average of 1,411 vs 18,407) in this sample than in the sample used in Table 2. Despite this, the covariance shrinkage method attributes 20 percent of the variation to bureaucrats and organizations.

[53]We are able to match 70 percent of the items assigned an 10-digit HS code in Step 2 of the text analysis method with the Sutton (1998) measure.

[54]Another possibility is that organizations endogenously respond to the effectiveness of bureaucrats by purchasing more/fewer, or different types of, goods. This would lead us to underestimate the true variance in procurer effectiveness.

[55]The algorithm developed in Step 2 of the procedure outlined in Sub-section II.A and Online Appendix A assigns a 10-digit code to 37 percent of the items in our analysis sample with high confidence.

These results reassure us both that our text analysis procedure accurately classifies purchases into homogenous categories and that our broad sample of products is appropriate.

Spending Quality

As discussed in Sub-section II.B, procurers' primary goal is to achieve low prices without sacrificing on item quality. However, prices are not the only procurement outcome that matters. We study a form of procurement where non-price goals are a priori less important than they are in services or public works contracts. Nevertheless, bureaucrats who procure off-the-shelf, manufactured goods may also face multitasking problems in balancing price against other objectives. If buyers who achieve low prices do poorly on other measures of performance, we may erroneously conclude that they are effective when a more comprehensive evaluation would not.

To investigate, we first repeat our analysis using our spending quality index as the outcome instead of prices. Table 4 shows the results. Two key findings emerge. First, time and product effects explain a far smaller share of the variation in spending quality than in prices. This is unsurprising insofar as production costs vary significantly across products. More interestingly, it does not appear that the contracting problems and delays captured by our spending quality index are concentrated among a subset of products, perhaps because we restrict our sample to a broad range of similarly homogeneous manufactured goods. Second, the four estimation methods from Table 2 reveal that a significant share of the variation in spending quality is driven by the procurers. The most conservative, covariance-shrinkage method, attributes 24 percent of the variation to bureaucrats and organizations. This is expected since the components of the spending quality index—particularly contract renegotiations and cost overruns—are outcomes buyers have scope to influence. However, bureaucrats and organizations explain a smaller share of the variation in spending quality than in prices.

To study the multitasking issue, Figure 3 shows the correlation between the bureaucrats' (Panel A) and organizations' (Panel B) covariance-shrunk price and spending quality effects. The panels show binned scatterplots together with a regression line fitted on the underlying data, and the correlation between the two effects is shown in the upper left corner. The figure reveals a strong, positive relationship between procurers' impact on the two outcomes and a fairly linear relationship between the two (correlations of 0.43 for bureaucrats and 0.48 for organizations): bureaucrats who achieve low prices also perform well on spending quality, and similarly for organizations. Additionally, in Appendix Table E.1 we re-estimate the variance decomposition including the spending quality outcomes as controls (despite them more properly being considered endogenous to the bu-

The remaining items in the Analysis Sample are also clustered into homogeneous bins, but we cannot confidently assign a pre-existing 10-digit code to these items.

reaucrat and organization making the purchase), and show that the results are essentially unchanged from our baseline specification in Table 2.

Overall, the results suggest that while bureaucrats and organizations clearly influence spending quality, the multitasking issue is not severe. In our subsequent analysis we thus use bureaucrats' and organizations' estimated effect on prices paid—their primary, legislated target (Yakovlev, Yakobson and Yudkevich, 2011)—as our preferred measure of their performance.

E What do effective bureaucracies do differently?

We now analyze what it is that distinguishes effective bureaucracies from their ineffective peers. Our data contain detailed information on the evolution of each of the 6.5 million procurement processes in the sample.[56] We construct 85 potential explanatory variables for bureaucrats, and 114 for organizations, which we summarize in Table F.1. There are seven categories of predictor variables: spending quality measures (6 variables for bureaucrats, 5 for organizations), and features of the purchase request (12); the bureaucrat and organization (12 and 42); the auction (19); participating suppliers (35); and the region (5). We investigate which of these co-vary with the estimated price- and spending quality-effectiveness of the implementing bureaucrat and organization.

To avoid overfitting and for the sake of parsimony, we use a LASSO procedure to first select 30 predictor variables.[57] We then regress each purchase's covariance-shrunk bureaucrat effect on these variables, the purchase's organization effect, and the controls in (3) (and vice-versa for the organization effects, the results for which are shown in appendix Figures F.1 – F.6).[58] The left panels of Figures 4 and 5 show coefficients from a series of bivariate regressions of the bureaucrat price effect (in Figure 4) and spending quality effect (in Figure 5) on each of the selected observables. The right panels show the LASSO coefficients (as crosses) and those from the multivariate regression of the procurer effect on all of the selected variables (as circles). To facilitate comparison, all variables are standardized to have unit standard deviation. The coefficients can thus be interpreted as the association between a one-standard deviation change in the predictor and the impact of the procurer on prices.

Several interesting patterns emerge. First, effective bureaucrats do not rush at the end of the fiscal year, a practice associated with wasteful spending (Lieb-

[56]In addition to the process, contract, firm, and spending quality data described in Section I, we here also use data on corruption and other measures of institutions across regions from Schulze, Sjahrir and Zakharov (2016).

[57]The procedure selects the smallest model with *at least* 30 predictors so the actual number varies slightly from figure to figure. Table F.1 shows pairwise coefficients from regressing price-effectiveness on each of the 411 potential explanatory variables we start out with. Figures F.3 and F.4 instead show results from using the LASSO procedure to select 60 instead of 30 predictors. The patterns in the findings are very similar to those described here.

[58]To account for small firms not being covered by the *Ruslana* data and the strong correlation between some of our variables, we also use an elastic net regularizer (a weighted average of LASSO and Ridge regression). Figures F.7 and F.8 show that the results are not sensitive to placing more weight on the Ridge regression.

man and Mahoney, 2017), and they complete documentation correctly, specifying product names and codes correctly more often. They also tend to make larger, more diverse, purchases, budling together several products. Ultimately, effective bureaucrats attract and admit a larger and more diverse pool of bidders, as emphasized by the theoretical framework in Section III. For each auction, we calculate the fraction of the pool of potential bidders who participate, and for each bureaucrat we calculate the Herfindahl-Hirschmann Index (HHI) of the suppliers they work with, and we find that both measures are strongly correlated with bureaucrat performance.[59] An example comes from a purchase of winter boots for a Saratov orphanage. The bureaucrat overseeing the request disqualified a firm from participating in the subsequent auction on the grounds that its application did not contain information on the height of the firm's boots' sole and heel. Only two bids were ultimately submitted in the auction, and the orphanage ended up paying a price per boot less than 10 percent below the maximum price.

Second, effectiveness appears to be very embodied in the individual procurers doing the work. Of the four categories of predictors we consider—features of the auction request; the bureaucrat; the auction itself; and the participating suppliers—characteristics of the bureaucrat have most predictive power. More experienced bureaucrats—for example those who run more auctions—are more effective, consistent with them having a larger network of contacts with suppliers to draw on. They also have fewer procurement processes fail due to no suppliers applying to participate.[60] Finally, bureaucrats who specialize more in particular products (as measured by the HHI of the products they buy)—another measure of bureaucrat experience—are also more effective.

Third, effective bureaucrats also end up purchasing from particular types of suppliers. They buy from suppliers that specialize in the products requested, and in selling to government (as measured by contracts won from state-owned enterprises). Effective bureaucrats also avoid middlemen: they are less likely to buy from wholesalers and exporters, but more likely to buy from firms that import the product they are purchasing. Finally, their suppliers are less likely to have the same postal code, or even to come from the same region.

Fourth, the correlates of bureaucratic effectiveness are strikingly similar when we look at prices paid and spending quality. 22 of the 34 strongest predictors of price effectiveness included in Figure 4 are also among the 32 strongest predictors of spending quality effectiveness in Figure 5.[61] This is particularly true for features of the bureaucrats themselves, where all the features selected for the price outcome also appear in the quality figure. This is not surprising since we saw in Sub-section IV.D that bureaucrats' price effectiveness is highly positively corre-

[59] We treat all winners of contracts for the same 2-digit product in the previous semester as the pool of potential bidders.

[60] We label the fraction of the bureaucrat's purchases where this does not occur their "success rate". Purchase failure is an uncommon outcome, but effectiveness and success being positively correlated also assuages a potential selection concern about only observing successful purchases in our main dataset.

[61] 20 of these have the same sign in both cases.

lated with their spending quality effectiveness. A final observation worth making is that there are some notable variables among those that are *not* selected by the LASSO. In particular, the wide range of regional measures of corruption have very weak predictive power. It thus appears that variation in bureaucratic procurement effectiveness in Russia is not primarily due to variation in corruption.

We conclude from these findings that a key part of what makes procurers effective is their ability to reduce entry barriers to participation in procurement auctions, and to attract firms to their auctions. The findings are very similar when we look at the determinants of organization effectiveness and so for conciseness we relegate those figures to Appendix F. One interesting additional finding that does emerge is that the single strongest predictor of an organization's effectiveness is its overall performance score in independent surveys and evaluations conducted by the Federal Treasury. This suggests both that our measure of effectiveness is correlated with what the federal government considers to be important, and that this effectiveness could be measured independently by central governments and then used to set procurement policy, the subject we turn to next.

V Policy Design with a Heterogeneous Bureaucracy

We saw in Section IV that a large share of the overall variation in performance under a constant policy regime is attributable to bureaucratic agents' effectiveness. But in many organizations—especially in the public sector—increasing productivity directly, through human resource practices, can be infeasible or costly. Such enterprises can instead change their task assignment, better tailoring work protocols to their workforce. In this section we study the introduction of a different policy regime in Russian procurement—a change in the bureaucracy's tasks. We show that the introduction of bid preferences favoring local manufacturers successfully shifted contracts to domestic producers, without significant impacts on prices or spending quality overall. However, these average treatment effects mask dramatic heterogeneity across "good" versus "bad" procurers, suggesting that there is significant scope for tailoring policy design to the effectiveness of the implementing bureaucracy.

A Overall impact of bid preferences for locally manufactured goods

Many governments use bid preferences to attempt to steer demand towards favored firms. The impact of such policies is theoretically ambiguous (see e.g. McAfee and McMillan, 1989), though empirical studies in contexts with high state capacity tend to find price increases and participation decreases (Marion, 2007; Krasnokutskaya and Seim, 2011; Athey, Coey and Levin, 2013). In Russia's case, as in many others, bid preferences favor local manufacturers. Its preferences policy imposed a bid penalty of 15 percent on foreign-manufactured goods (see Sub-section I.C). In 2011–2014, the preferences only came into effect in May or June each year. Moreover, the policy applied only to a subset of goods—a subset that varied from year to year.[62] We exploit this variation in a generalized difference-in-differences design, estimating

$$y_{igt} = \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \delta \text{Preferenced}_{gt} \times \text{PolicyActive}_t + \varepsilon_{igt} \tag{5}$$

where y_{igt} is the outcome in purchase i of good g in month t, Preferenced_{gt} is a dummy indicating that g is a treated good in the year month t falls within, and PolicyActive_t is a dummy indicating that the year's list of preferenced goods has been published. $\mathbf{X}_{igt}$ are the same controls we use in Section IV, but for clarity we separate out the good and month fixed effects, μ_g and λ_t. ε_{igt} is an error term we allow to be clustered by month and good. Because there must be a minimum of one bidder in the auction offering a Russian-made good and a minimum of one bidder offering a foreign-made good for preferences to apply, our estimates should be interpreted as Intent-to-Treat (ITT) effects.

Following Cengiz et al. (2019), we also stack all the events (the preference list being published) to estimate an event study analog of equation (5) in a window starting three months before and ending four months after each year's preference list is published (ListMonth_s):

$$p_{igt} = \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \sum_{s=-3}^{4} \delta_s \text{Preferenced}_{gt} \times \mathbf{1}\left\{t - \text{ListMonth}_t = s\right\} + \varepsilon_{igt} \tag{6}$$

To estimate (5) and (6), we expand the Analysis Sample and Pharmaceuticals Sub-sample to also include purchases where bid preferences apply, and which were managed by bureaucrats and organizations in these samples. The samples are summarized in columns (3) and (6) of Table 1. In the Analysis Sample, we define Preferenced_{gt} as a dummy equal to one if good g is on that year's list. Since pharmaceuticals are always on the list, for pharmaceuticals we instead define Preferenced_{gt} as equal to one if the drug is manufactured both in Russia and abroad.[63]

The estimated event study coefficients δ_s are all close to zero and statistically indistinguishable from zero in the months leading up to the publication of the preference list. Figure 6 shows this for prices in the Analysis Sample. This finding lends credibility to our difference-in-differences design's identifying assumption of parallel trends. The figure also shows no evidence of anticipation of the publication of the preference list. Figure G.1 shows the evolution of the share of purchases for preferenced items around the date of the publication of the list and also shows no evidence that buyers are able to manipulate the timing of their purchases to avoid or take advantage of preferences.

[62] Preferenced goods spanned many categories, including automobiles, clocks, various food products, medical equipment, pharmaceuticals, and textile and furs (see Table H.1 for the full list).

[63] Several drugs in use in Russia are manufactured either only abroad or only domestically.

The preferences policy achieves its primary goal: the good purchased is 14 percent more likely to be domestically manufactured when bid preferences are in effect.[64] We show this first result from estimating (5) in Column (7) of Table 5. Columns (1) to (6) establish that it does so at little to no cost. Participation declines somewhat in both the full sample and the pharmaceuticals sample. However, prices are unaffected on average and spending quality increases slightly in the full sample; in the pharmaceuticals sample prices decrease and spending quality decreases somewhat on average.[65] The limited or even beneficial overall impact on prices suggests that the policy's discouragement of foreign manufacturers is offset by a combination of encouragement of local manufacturers and the mechanical decrease that applies when the winning bidder supplies foreign manufactured goods.

These findings contrast with studies of similar preference policies in the U.S. (see e.g. Marion, 2007; Krasnokutskaya and Seim, 2011; Athey, Coey and Levin, 2013). Our analysis in the next sub-section points towards a possible explanation: U.S. procurers are probably more effective on average than Russian procurers. We estimate impacts similar to those found in the U.S.—increased prices—when preferences are implemented by Russian procurers of high effectiveness, but when procurers are ineffective, we find the opposite impact.

B *Bureaucratic performance heterogeneity under different policy regimes*

The model in Section III implies that bid preferences will compress the procurement performance of the bureaucratic apparatus. Proposition 2 describes how the variation in the entry costs buyers impose on suppliers that drives bureaucracies' effectiveness can also lead to patterns of heterogeneity in the treatment effect of introducing bid preferences. Such a finding would have striking implications for procurement policy design across contexts. To test this proposition in our data, we now compare treatment effects among effective and ineffective buyers. Estimates of effectiveness (in the absence of bid preferences) come from our analysis in Section IV.

We extend (5) to estimate heterogeneous treatment effects as follows:

$$\begin{aligned} y_{igt} = \mathbf{X}_{igt}\beta + \mu_g + \lambda_t + \theta_b\hat{\alpha}_b + \theta_j\hat{\psi}_j + \delta\text{Preferenced}_{gt} \times \text{PolicyActive}_t + \rho_b\text{Preferenced}_{gt}\hat{\alpha}_b \\ + \rho_j\text{Preferenced}_{gt}\hat{\psi}_j + \eta_b\text{PolicyActive}_t\hat{\alpha}_b + \eta_j\text{PolicyActive}_t\hat{\psi}_j \\ + \pi_b\text{Preferenced}_{gt} \times \text{PolicyActive}_t\hat{\alpha}_b + \pi_j\text{Preferenced}_{gt} \times \text{PolicyActive}_t\hat{\psi}_j + \varepsilon_{igt} \end{aligned} \tag{7}$$

Table 6 shows the results. The small negative average price effect from Sub-

[64] In Column (7) of Table 5 we restrict the sample to purchases in which an auction takes place in order to be consistent with Column (7) of Table 6. We find an increase in the probability of a domestic producer winning the auction of similar magnitude in the full pharmaceuticals sample (results available from the authors upon request).

[65] Recall that a higher number implies worse spending quality.

section V.A masks substantial heterogeneity in the impact of bid preferences across bureaucracies. Consistent with Proposition 2, we find that prices drop significantly more for bureaucrats who pay higher prices when there are no bid preferences (i.e., who have a higher $\hat{\alpha}_b$). Columns (1) and (4) of Table 6 show this stark pattern both in the full sample and the pharmaceuticals sample. As we return to in Sub-section V.D, the estimated coefficient on Bureaucrat FE $\times$ Preferenced $\times$ Policy Active is large (in absolute value) in both samples, and especially so in the pharmaceuticals sample.

Consistent with the model, these price improvements are accompanied by increases in participation (columns (2) and (5) of Table 6) and do not come at the expense of spending quality (columns (3) and (6)). Column (7) shows that the increases in the probability of a domestic winner are also concentrated among the least effective bureaucrats.[66]

While we find support for all the model's predictions on heterogeneous effects of bid preferences by bureaucrat effectiveness, we do not see this for organization effectiveness $\hat{\psi}_j$.[67] The model offers a potential explanation: heterogeneity of the effect of bid preferences is driven by participation costs, but differences in organization effectiveness may to a greater extent be due to contract fulfillment costs than participation costs.[68]

We next estimate a less parametric version of (7) by including separate triple-interaction terms for each decile of bureaucrat effectiveness $\hat{\alpha}_b$ and organization

[66]In fact, Proposition 2 predicts a U-shaped relationship between the probability a domestic good is supplier and bureaucrat type. Panel B of Figure G.4 shows that this is indeed what we see. We do not see a similar pattern for organizations and in fact the negative coefficient in Column (7) is not picking up a strong pattern of smaller effects for less effective organizations (results available upon request).

[67]That is, we see price decreases that are largest for the least effective bureaucrats; changes in participation that are larger for the least effective bureaucrats; and a U-shaped relationship between the probability a domestic good is supplied and bureaucrat type (results available upon request). When we look at heterogeneity by organization effectiveness, we do not see evidence consistent with any of these predictions. In the pharmaceuticals sample the coefficient on Organization FE $\times$ Preferenced $\times$ Policy Active is in fact positive, but very imprecisely estimated. In the full sample the differential effect for effective organizations is positive and marginally significant, but small in magnitude and in particular much smaller than the opposite-signed effect for effective bureaucrats.

[68] In the model, buyers impose two types of costs on potential suppliers: fulfillment costs (α_θ for bureaucrats and ψ_θ for organizations) and participation costs (α_c and ψ_c). As Proposition 1 states, both costs affect prices at baseline (without bid preferences) in the same way (though with different coefficients on fulfillment and participation costs) and so we subsumed them into the composite terms $\tilde{\alpha}$ and $\tilde{\psi}$ that enter equation (1). By contrast, as Proposition 2 states, the heterogeneity of the effects of bid preferences is governed by the participation costs and not the fulfillment costs. Hence, if most of the variation in baseline performance of organizations is driven by fulfillment costs, while most of the variation in baseline performance of bureaucrats is driven by participation costs, then we would expect Proposition 2 to be consistent with the heterogeneity of the estimated treatment effects by bureaucrat effectiveness but not by organization effectiveness, which is what we see. Consistent with this, when we compare the features that predict baseline effectiveness for bureaucrats vs organizations, we see differences. The predictors of organization effectiveness are less to do with participation (the types of costs we think may be incorporated in α_c and ψ_c) than those for bureaucrat effectiveness. Those that *do* predict organization effectiveness are more to do with the end user organization itself and, potentially, their idiosyncratic fulfillment costs.

effectiveness $\hat{\psi}_j$:

$$y_{igt} = \sum_{k=1}^{10} \{D_{kj} + D_{kb} \times (\rho_k \text{Preferenced}_{gt} + \eta_k \text{PolicyActive}_t + \pi_k \text{Preferenced}_{gt} \times \text{PolicyActive}_t)\}$$
$$+ \mathbf{X}_{igt}\beta + \mu_g + \mu_t + \varepsilon_{igt} \quad (8)$$

where D_{kj} and D_{kb} are indicators for organization j and bureaucrat b belonging to decile k of their respective distributions of effectiveness. We also extend the event study (6) to estimate effects separately by quartile of bureaucrat- and organization-effectiveness. In these, rather than normalizing the reference month (the month before the preference list is published) to zero, we normalize it to the baseline performance in each group to better highlight how different their performance was before the preferences are introduced, and how their performance converges as a result of the preferences.

The price decreases in Table 6 are concentrated among the least effective bureaucrats. Figure 7 shows this graphically. We see a clear pattern of larger price drops for ineffective bureaucrats in Panel A, with the estimated price effect of the policy decreasing throughout the lowest deciles of bureaucrat effectiveness. The figure also shows more suggestive evidence of price *increases* when the policy is administered by effective bureaucrats.[69] The event studies in Panel B of Figure 7 help rule out potential confounds like mean reversion or differences in seasonality across different types of bureaucrats. The graph shows no discernible trends in prices before the introduction of bid preferences and then a marked divergence of prices paid by the two groups—high versus low effectiveness bureaucrats—after the introduction of preferences. These patterns provide compelling evidence that the estimates in Table 6 capture the causal differential of interest.

Overall, these results suggest that, from the perspective of a government trying to minimize the prices it pays for its goods while simultaneously steering government demand towards domestic manufacturers, a "buy local" procurement policy of the form used in Russia is a more effective policy tool when the bureaucrats administering the policy are *less* effective at their job, consistent with the logic of our model in Section III. We trace out the policy design implications in Sub-section V.D, after examining what explains this heterogeneity in policy impact in Sub-section V.C.

[69] Appendix Figure G.2 shows the analogous results for organizations, confirming the findings in table 6 that there is limited heterogeneity. Appendix Figure G.3 shows that consistent with the findings for prices, we see strong heterogeneity in the impact of the policy on participation by bureaucrat effectiveness. More effective bureaucrats experience large drops in participation, while less effective bureaucrats do not experience these participation drops, and may even see participation increases. Similarly to the effects on prices, there is little evidence of heterogeneity by organization effectiveness.

C Drivers of performance heterogeneity under different policy regimes

To unpack the relationship between bureaucratic heterogeneity and performance under different policy regimes, we turn again to our data on procurement processes and take an approach similar to the one we used in Sub-section IV.E to study the drivers of performance in the baseline policy regime. Bureaucratic effectiveness can affect policy performance under different policy regimes in two ways. First, the attributes that are associated with bureaucratic effectiveness in the baseline policy regime may assume a different significance under the preference policy regime even without the bureaucrats or organizations changing the way they carry out their work. Second, new attributes may become important under the preference regime, and so bureaucrats and organizations that are able to change these attributes may benefit the most from the policy change.

We estimate a triple difference regression akin to (7):

$$\begin{aligned} y_{igt} = {} & \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \mathbf{Z}_{igt}\boldsymbol{\theta} + \text{Preferenced}_{gt} \times \mathbf{Z}_{igt}\boldsymbol{\gamma} + \text{PolicyActive}_t \times \mathbf{Z}_{igt}\boldsymbol{\eta} \\ & + \delta\text{Preferenced}_{gt} \times \text{PolicyActive}_t + \text{Preferenced}_{gt} \times \text{PolicyActive}_t \times \mathbf{Z}_{igt}\boldsymbol{\pi} + \varepsilon_{igt} \end{aligned} \tag{9}$$

where all terms are as defined in equation (7) but we replace the interactions with bureaucrat ($\hat{\alpha}_b$) and organization ($\hat{\psi}_j$) effectiveness with a vector of observables $\mathbf{Z}_{igt}$. Since our data contain a large number of these (see Table F.1), the vector $\mathbf{Z}_{igt}$ is chosen by the same regularization procedure used in Sub-section IV.E.[70] Comparing the variables in the vector $\mathbf{Z}_{igt}$ selected here to those selected when studying the correlates of baseline performance in Sub-section IV.E allows us to tell apart the two channels discussed above.

Figure 8 shows the results for prices paid. The variables that affect the policy's impact without changes in bureaucratic behavior are those relating to the bureaucrats themselves (3 out of 5 variables in Figure 4 are also in $\mathbf{Z}_{igt}$), and somewhat the organizations (3/22 variables). These variables describe the buyers rather than how the auction plays out (1 of 8 variables) and who the eventual supplier is (4 of 13 variables). Without changes in behavior, these upstream factors influence how the policy change affects procurement performance more than downstream factors like the types of participants and how auctions play out. Conversely, the variables that become important under preferences are those relating to the suppliers (4/13 variables). Particularly noteworthy is that the share of bidders in the auctions who have experience importing or have foreign ownership become relevant, presumably since the preference policy drives a wedge between foreign and domestic products.[71]

[70]We first run a LASSO procedure with the full set of observables in our data to select the elements of $\mathbf{Z}_{igt}$. For the selected variables, we run regression (9). As in Sub-section IV.E, we also use an elastic net procedure so that the regularization takes greater account of the correlation between the observables. Figures G.6 (for prices) and G.7 (for quality) show that the results are very robust to how much weight we place on the ridge criterion in the elastic net.

[71]Figure G.5 shows that the results are very similar when we study spending quality as the outcome

Summarizing, the same upstream characteristics of the buyers and the way they write requests drive baseline performance and the impacts of policy change. But under the new policy, different characteristics of the auction and the supplier matter. Ultimately, this suggests that there is significant scope for tailoring policy design to the capacity of the implementing bureaucracy since it is these deeper characteristics embodied in the buyers that appear to matter under both policy regimes. These results illuminate *why* the potential scope for and benefits of tailoring policy design to the capacity of implementing bureaucrats are as large as the results in Sub-section V.B suggest.

D Implications for policy design

We have seen that deviations from mechanistic, uniform performance—Weber's ideal—depend not just on a bureaucracy's workforce, but also on the policies that these individuals and organizations are asked to carry out. The model in Section III illustrates why commonly observed preferences for domestic producers may plausibly achieve public procurement goals in polities with ineffective bureaucracies, but not in polities with effective bureaucracies. We found evidence that this is in fact the case for a 15 percent preference rate in Figure 7A: the policy decreased prices by up to 14 percent when implemented by the least effective Russian bureaucrats, but for more effective bureaucrats, prices *increased*. The adverse impact when the policy is administered by effective bureaucrats in Russia is comparable to that for similar preference policies in the U.S. (Marion, 2007; Krasnokutskaya and Seim, 2011; Athey, Coey and Levin, 2013). This raises the question of whether policy makers may want to pick different bid preference policies depending on the effectiveness of their implementing bureaucracy.

To shed light on this question, we again need both the bureaucrat effectiveness estimates from Section IV and the heterogeneity-in-impact estimates from Subsection V.B. We combine our estimates from Figure 7 of the effect of the 15 percent preference in each decile of the overall effectiveness distribution with the distributions of effectiveness in a range of subgroups of bureaucrats. Assuming that the semi-elasticity of prices with respect to the preference rate is locally constant, we can then estimate the level of the preference rate that would achieve the same effect in each subgroup as we observe on average across the full sample.

Specifically, we assume that for each decile k of effectiveness, the semi-elasticity of prices with respect to the preference rate $1-\gamma$ is equal to the average treatment effect of the 15 percent preference rate we estimate for that decile so that log prices are locally linear in the preference rate, with slope $TE_k/0.15$, where TE_k is the treatment effect for decile k estimated using equation (8) shown in Figure 7A. This is a strong assumption, and the model in Section III does not imply this constant elasticity, but we show in Appendix Figure G.8 that such a simplification is nevertheless reasonable locally.

instead of prices.

We can then ask, for any subgroup g with distribution of bureaucrats w_{kg}, $k = 1, \ldots, 10$ across the deciles of effectiveness, what preference rate $1 - \gamma_g^*$ would achieve the same impact in that subgroup as the 15 percent rate achieves in the overall sample. For each subgroup, our estimates in Figure 7A imply a treatment effect of a 15 percent bid penalty of $TE_g = \sum_{k=1}^{10} w_{kg} TE_k$, and our constant elasticity assumption implies that we can find the equivalent policy by solving

$$d \log (p_g) = \overline{TE} - TE_g = \left(1 - \gamma_g^* - 0.15\right) \frac{TE_g}{0.15} \Leftrightarrow 1 - \gamma_g^* = \frac{0.15\overline{TE}}{TE_g} \tag{10}$$

where $\overline{TE} = \sum_{k=1}^{10} TE_k$ is the treatment effect in the overall sample. Applying equation (10) in different subgroups allows us to provide a back of the envelope estimate of how policy-makers overseeing different bureaucracies can achieve a given policy goal, in this case a particular overall effect on prices, by tailoring the preference policy to the effectiveness of the implementing bureaucracy.

We consider subgroups of bureaucrats distinguished by the government department they are working with, the level of government they work with, their experience (the volume of transactions they undertake), and whether they work in-house or externally. These are observable markers that we consider in our analysis of the drivers of bureaucratic performance, and which policy-makers might plausibly consider when designing policy. Figure 9 plots these subgroups' equivalent bid penalty $1 - \gamma_g^*$ from equation (10) against the group's average baseline performance, excluding groups for which the 95% confidence interval on $1 - \gamma_g^*$ is wider than 0.3.[72]

Figure 9 shows a wide range of equivalent policies, ranging from 23 percent for the most effective subgroup, to 10 percent for the least effective subgroup. These numbers are, of course, obtained under highly restrictive assumptions, but they nevertheless serve to illustrate the usefulness of considering individual policy-implementers' effectiveness in policy design.

VI Conclusion

In this paper we have presented evidence that, contrary to the mechanistic view of the bureaucracy in much of the existing literature, the individuals and organizations tasked with implementing policy are important sources of variation in states' productivity. Bureaucrats and public sector organizations together account for a full 39 percent of the variation in quality-adjusted prices paid by the Russian government for its inputs. Consistent with a simple endogenous entry model of procurement, effective public procurers engage in practices that lower entry costs for potential suppliers and attract a larger and more diverse pool of participants, allowing them to achieve lower prices. However, in many contexts,

[72]These tend to be groups with very good baseline performance with many bureaucrats in deciles with estimated treatment effects vey close to zero, leading to noisy estimates when we divide through by them.

the performance of individuals and organizations cannot be directly improved, but the tasks bureaucrats are directed to carry out can. Studying the impact of a "buy local" policy that provides bid preferences for locally manufactured goods, we show that participation increases and prices decrease when the policy is implemented by less effective bureaucrats, while performance is essentially unaffected when the policy is implemented by more effective bureaucrats, consistent with our model.

These findings have important implications. First, they suggest that there are huge returns to the state from employing more bureaucrats at the high end of the observed performance range, training bureaucrats better, or improving organization-wide characteristics such as management quality—if such changes are possible. Second, our findings imply that the nature of the policy regime in place determines the extent to which differences in bureaucratic effectiveness manifest themselves in differences in public sector output. In turn, this suggests that policies that are suboptimal when state effectiveness is high may become second-best optimal when state effectiveness is low. Achieving the *best* policy outcomes likely requires both improving the effectiveness of the bureaucratic apparatus and choosing policies that are tailored to the effectiveness of their implementers.

References

Abowd, John M., Francis Kramarz, and David Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67: 251–333.

Abowd, John M., Robert H. Creecy, and Francis Kramarz. 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Census Bureau Technical Paper TP-2002-06.

Allcott, Hunt. 2015. "Site Selection Bias in Program Evaluation." *Quarterly Journal of Economics*, 130: 1117–1165.

Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Richard Upward. 2008. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias." *Journal of the Royal Statistical Society*, 171(3): 673–697.

Angrist, Joshua D. 2004. "Treatment effect heterogeneity in theory and practice." *The Economic Journal*, C52–C83.

Athey, Susan, Dominic Coey, and Jonathan Levin. 2013. "Set-Asides and Subsidies in Auctions." *American Economic Journal: Microeconomics.*

Bandiera, Oriana, Andrea Prat, and Tommaso Valletti. 2009. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review*, 99(4): 1278–1308.

Barabashev, Alexei, and Jeffrey D. Straussman. 2007. “Public service reform in Russia, 1991–2006.” *Public Administration Review*, 67(3): 373–382.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott. 2007. “Firms in international trade.” *Journal of Economic perspectives*, 21(3): 105–130.

Bertrand, Marianne, and Antoinette Schoar. 2003. “Managing with style: The effect of managers on firm policies.” *Quarterly Journal of Economics.*

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu. 2020. “The Glittering Prizes: Career Incentives and Bureaucrat Performance.” *Review of Economic Studies*, 87: 626–655.

Besley, Timothy, and Torsten Persson. 2009. “The origins of state capacity: Property rights, taxation, and politics.” *American Economic Review*, 99: 1218–44.

Best, Michael Carlos, Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. “Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan.” *Journal of Political Economy*, 123: 1311–1355.

Bloom, Nicholas, Jae Song, David Price, Fatih Guvenen, and Till von Wachter. 2019. “Firming up inequality.” *Quarterly Journal of Economics*, 134(1): 1–50.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng’ang’a, and Justin Sandefur. 2018. “Experimental Evidence on Scaling Up Education Reforms in Kenya.” *Journal of Public Economics*, 168: 1–20.

Bronnenberg, Bart, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M. Shapiro. 2015. “Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium.” *Quarterly Journal of Economics.*

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. “Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women.” *Quarterly Journal of Economics.*

Card, David, Ana Rute Cardoso, Jörg Heining, and Patrick Kline. 2018. “Firms and Labor Market Inequality: Evidence and Some Theory.” *Journal of Labor Economics*, 36: s13–s70.

Card, David, Jörg Heining, and Patrick Kline. 2013. “Workplace Heterogeneity and the Rise of West German Wage Inequality.” *Quarterly Journal of Economics*, 128: 967–1015.

Carneiro, Anabela, Paulo Guimarães, and Pedro Portugal. 2012. "Real Wages and the Business Cycle: Accounting for Worker, Firm and Job Title Heterogeneity." *American Economic Journal: Macroeconomics*, 4: 133–152.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs." *Quarterly Journal of Economics*, 134: 1405–1454.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014*a*. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104: 2593–2632.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014*b*. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review*, 104: 2533–2679.

ClearSpending. 2022. "Anomalies Data." https://clearspending.ru/in-control/.

Coviello, Decio, Andrea Guglielmo, and Giancarlo Spagnolo. 2018. "The Effect of Discretion on Procurement Performance." *Management Science*.

Coviello, Decio, Luigi Moretti, Giancarlo Spagnolo, and Paola Valbonesi. 2017. "Court Efficiency and Procurement Performance." *Scandinavian Journal of Economics*.

Dal Bo, Ernesto, Fred Finan, and Martin Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics*, , (128).

Decarolis, Francesco, Leonardo M. Giuffrida, Elisabetta Iossa, Vincenzo Mollisi, and Giancarlo Spagnolo. 2018. "Bureaucratic Competence and Procurement Outcomes." Mimeo Universita Bocconi.

Dehejia, Rajeev, Cristian Pop-Eleches, and Cyrus Samii. forthcoming. "From Local to Global: External Validity in a Fertility Natural Experiment." *Journal of Business and Economic Statistics*.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India." *Quarterly Journal of Economics*, 128: 1499–1545.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *Econometrica*, 86(6): 2123–2160.

EIS. 2022. "Unified Procurement Information System." Open Data.

Enikolopov, Ruben, and Ekaterina Zhuravskaya. 2007. “Decentralization and political institutions.” *Journal of Public Economics*.

Ferraz, Claudio, Frederico Finan, and Dimitri Szerman. 2015. “Procuring Firm Growth: The Effects of Government Purchases on Firm Dynamics.” mimeo, UC Berkeley.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande. 2017. “The Personnel Economics of the Developing State.” In *Handbook of Field Experiments, Volume II..* , ed. Abhijit Banerjee and Esther Duflo. North Holland.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. “Sources of Geographic Variation in Health Care: Evidence from Patient Migration.” *Quarterly Journal of Economics*.

Gaure, Simen. 2013. “OLS with multiple high dimensional category variables.” *Computational Statistics and Data Analysis*, 66: 8–18.

Gentry, Matthew, and Tong Li. 2014. “Identification in auctions with selective entry.” *Econometrica*, 82(1): 315–344.

Goldberg, Pinelopi Koujianou, and Jan De Loecker. 2014. “Firm Performance in a Global Market.” *Annual Review of Economics*.

Hansman, Christopher, Jonas Hjort, and Gianmarco León. 2019. “Interlinked Firms and the Consequences of Piecemeal Regulation.” *Journal of the European Economic Association*.

Heckman, James, and Jeffrey Smith. 1997. “Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts.” *Review of Economic Studies*, 487–535.

Hoberg, Gerard, and Gordon Phillips. 2016. “Text-Based Network Industries and Endogenous Product Differentiation.” *Journal of Political Economy*, 124(5): 1423–1465.

Jones, Benjamin F., and Benjamin A. Olken. 2005. “Do leaders matter? National leadership and growth since World War II.” *Quarterly Journal of Economics*.

Kane, Thomas J., and Douglas O. Staiger. 2008. “Estimating teacher impacts on student achievement: An experimental evaluation.” NBER working Paper.

Khan, Adnan, Asim Khwaja, and Ben Olken. 2016. “Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors.” *Quarterly Journal of Economics*.

Khan, Adnan, Asim Khwaja, and Ben Olken. 2018. "Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings." *American Economic Review.*

Khandelwal, Amit. 2010. "The Long and Short (of) Quality Ladders." *The Review of Economic Studies*, 77(4): 1450–1476.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica*, 75: 83–119.

Krasnokutskaya, Elena, and Katja Seim. 2011. "Bid Preference Programs and Participation in Highway Procurement Auctions." *American Economic Review*, , (101): 2653–2686.

Lacetera, Nicola, Bradley Larsen, Deving G. Pope, and Justin Sydnor. 2016. "Bid Takers or Market Makers? The Effect of Auctioneers on Auction Outcomes." *American Economic Journal: Microeconomics*, 8: 195–229.

Laffont, Jean-Jaques. 2005. *Regulation and Development.* Cambridge University Press.

Lancaster, Tony. 2000. "The Incidental Parameter Problem since 1948." *Journal of Econometrics*, 95: 391–413.

Lewis-Faupel, Sean, Yusuf Neggers, Benjamin Olken, and Rohini Pande. 2016. "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia." American Economic Journal: Economic Policy.

Liebman, Jeffrey B., and Neale Mahoney. 2017. "Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement." *American Economic Review*, 107: 3510–3549.

Marion, Justin. 2007. "Are Bid Preferences Benign? The Effect of Small Business Subsidies in Highway Procurement Auctions." *Journal of Public Economics*, , (91): 1591–1624.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review.*

McAfee, Preston R., and John McMillan. 1989. "Government procurement and international trade." *Journal of International Economics*, 26(3): 291–308.

Milgrom, Paul. 2004. *Putting Auction Theory to Work.* Cambridge: Cambridge University Press.

MinZrav. 2016. "List of Vital and Essential Drugs for Medical Use for 2016." Russian Ministry of Health.

Page, Lucy, and Rohini Pande. 2018. "Ending Global Poverty: Why Money Isn't Enough." *Journal of Economic Perspectives*, 32: 173–200.

Rasul, Imran, and Daniel Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *Economic Journal*.

Rauch, James E. 1999. "Networks Versus Markets in International Trade." *Journal of International Economics*, 7–35.

Rosstat. 2022*a*. "Accounting (Financial) Statements of Enterprises and Organizations." http://www.gks.ru/opendata/.

Rosstat. 2022*b*. "Russian State Statistics Agency." https://www.opendata.gov.ru/opendata/7708234640-urid-1.

Ruslana. 2022. "Bureau Van Dijk." Company Database.

Samuelson, William F. 1985. "Competitive bidding with entry costs." *Economics Letters*, 17(1): 53–57.

Schapper, Paul R., Joao Nuno Veiga Malta, and Diane L. Gilbert. 2009. "Analytical Framework for the Management and Reform of Public Procurement." In *International Handbook of Public Procurement.* , ed. Khi V. Thai. CRC Press.

Schulze, Günther G, Bambang Suharnoko Sjahrir, and Nikita Zakharov. 2016. "Corruption in Russia." *The Journal of Law and Economics*, 59(1): 135–171.

Silver, David. 2016. "Haste or Waste? Peer Pressure and the Distribution of Marginal Returns to Health Care." Mimeo: UC Berkeley.

Sutton, John. 1998. *Technology and Market Structure: Theory and History.* Cambridge: MIT Press.

Syverson, Chad. 2004. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy.*

Szakonyi, David. 2018. "Businesspeople in Elected Office: Identifying Private Benefits from Firm-Level Returns." *American Political Science Review*, 112(2): 322–338.

Transparency International. 2016. "Corruption Perceptions Index." 2016 version.

Weber, Max. 1921. *Economy and Society.*

Xu, Guo. 2018. "The Costs of Patronage: Evidence from the British Empire." *American Economic Review*, 108(11): 3170–3198.

Yakovlev, Andrei, Lev Yakobson, and Maria Yudkevich. 2011. "The Public Procurement System in Russia: Road Toward A New Quality." Unpublished Working Paper.

Yakovlev, Andrei, Olga Demidova, and Olga Balaeva. 2010. "The System of Public Procurements in Russia: On the Road of Reform." *HSE Policy Paper*, , (1): 1–23.

Yakovlev, Evgeny, and Ekaterina Zhuravskaya. 2014. "The Unequal Enforcement of Liberalization: Evidence from Russia's Reform of Business Regulation." *Journal of the European Economic Association*.

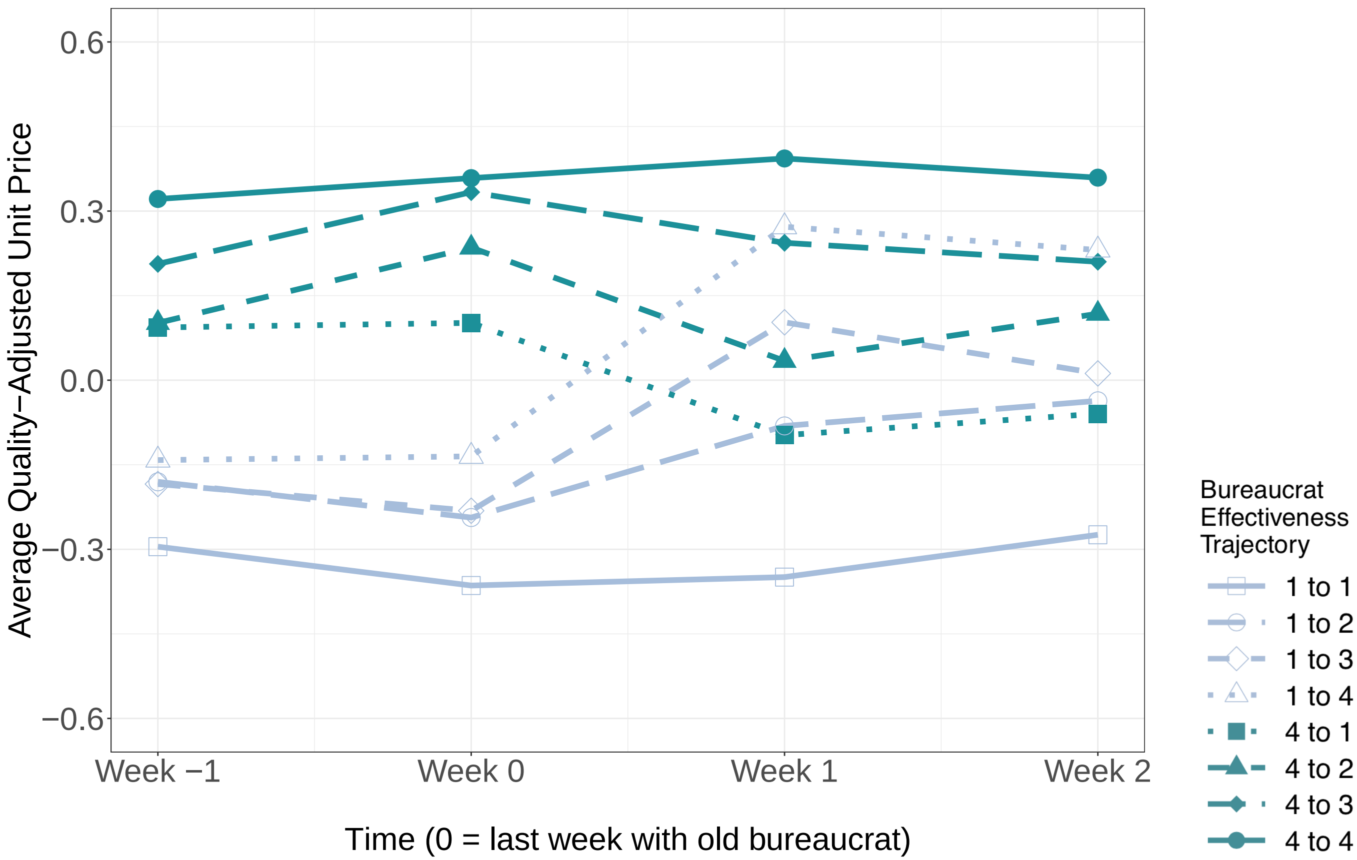


Figure 1. Event Study of Procurement Prices Around Times Organizations Switch Bureaucrats

Note: The figure shows time trends in prices around the time that organizations switch which bureaucrat makes purchases on their behalf. The horizontal axis indexes weeks in which bureaucrat-organization pairs work together, with time 0 being the last week in which the organization works with the old bureaucrat just before switch, and time 1 being the first week the organization works with the new bureaucrat after the switch. The y axis measures average residualized prices paid by the bureaucrat-organization pair where prices are residualized by regressing log unit prices on good and month fixed effects. We create a balanced panel in which we require each bureaucrat-organization pair to work together in at lesat two separate weeks and each bureaucrat to work with at least one other organization in the quarter containing time 0 (for the "old" bureaucrat the organization works with before the switch) or time 1 (for the "new" bureaucrat the organization works with after the switch). Bureaucrats are classified into quartiles according to the average (residualized) prices they achieve with the *other* organizations they work with in the quarter containing time 0 (for the old bureaucrat) or the quarter containing time 1 (for the new bureaucrat).

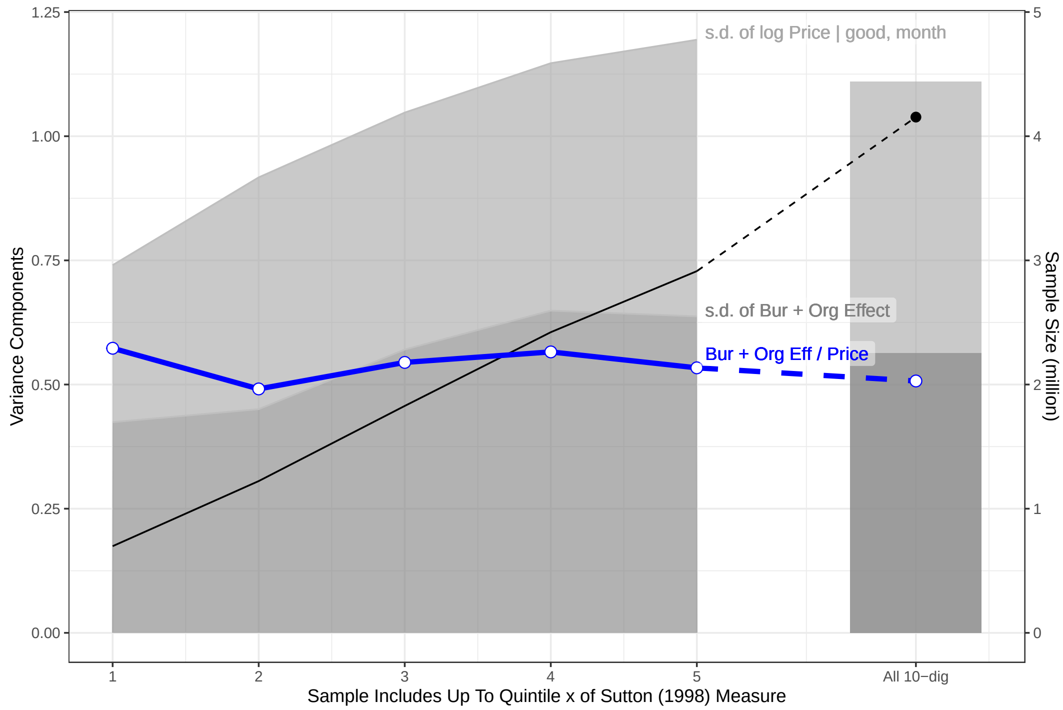


Figure 2. Robustness to Using Subsamples of Increasingly Heterogeneous Goods

Note: The figure shows the components of the variance of prices due to bureaucrats and organizations estimated by implementing the variance decomposition in equation (4) (see notes to Table 2 for details). The right-most bar uses the sub-sample consisting of all auctions for goods that our text analysis classification method is able to assign a 10-digit product code to. The left portion of the figure uses the sub-set of the sample that we can match to the scope-for-quality-differentiation ladder developed by Sutton (1998). Moving from right to left we remove quintiles of the data with the highest scope for quality differentiation, as shown by the black line, which indicates the sample size used. The dark shaded region is the variance of prices attributable to the bureaucrats and organizations. The dark and light shaded regions show the total variance of prices. The blue line shows the fraction of the overall variance attributable to bureaucrats and organization, highlighting that it remains roughly constant as we add more heterogeneous goods to the sample.

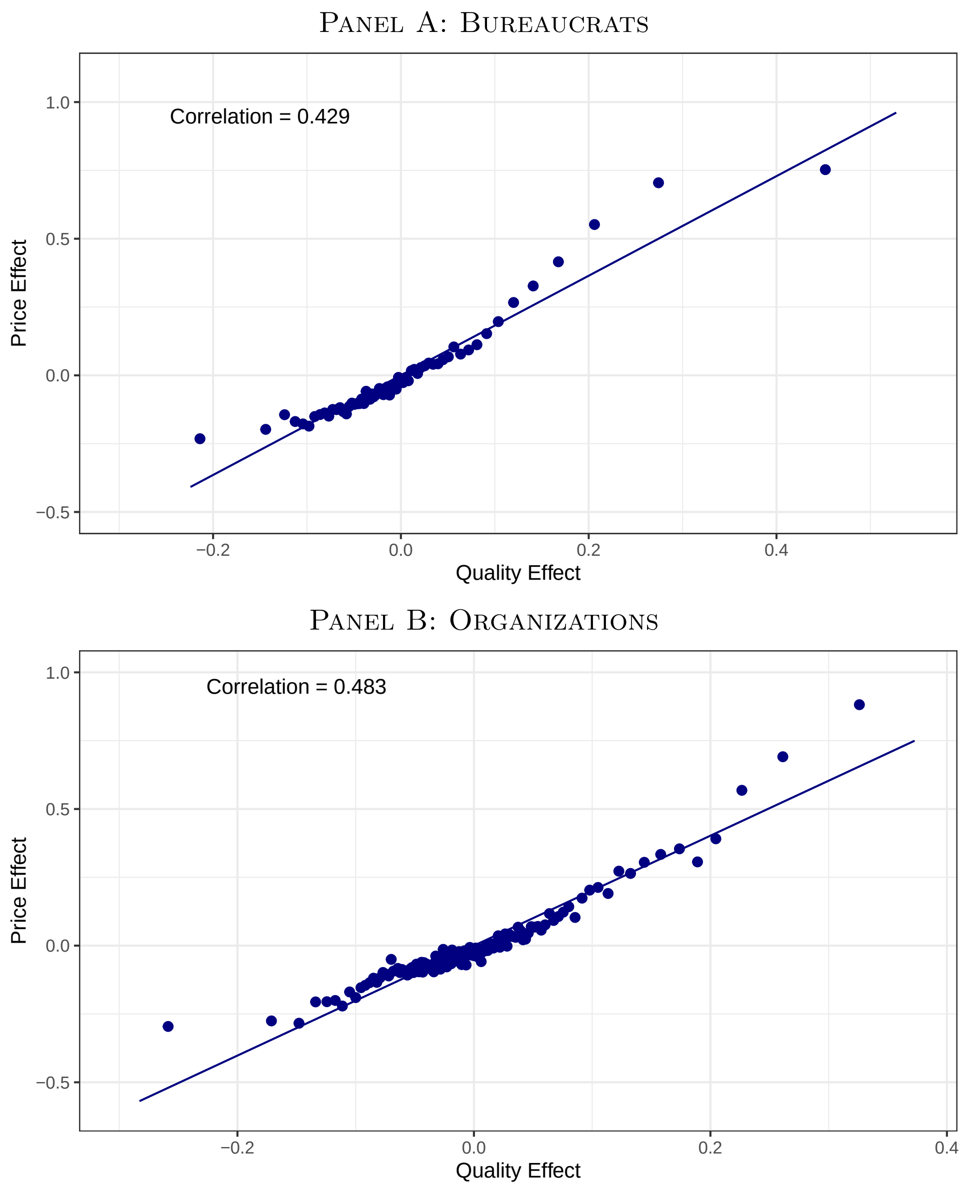


Figure 3. Buyers Who Achieve Low Prices Also Achieve Better Spending Quality

Note: The figure shows the correlation between bureaucrats' (panel A) and organizations' (panel B) covariance-shrunk price effects and their covariance-shrunk spending quality effects. They are estimated by implementing the variance decomposition in equation (4) and then implementing our covariance-shrinkage method (see notes to Table 2 for details). The panels show binned scatterplots together with a regression line fitted on the underlying data, and the correlation between the two effects shown in the upper left corner.

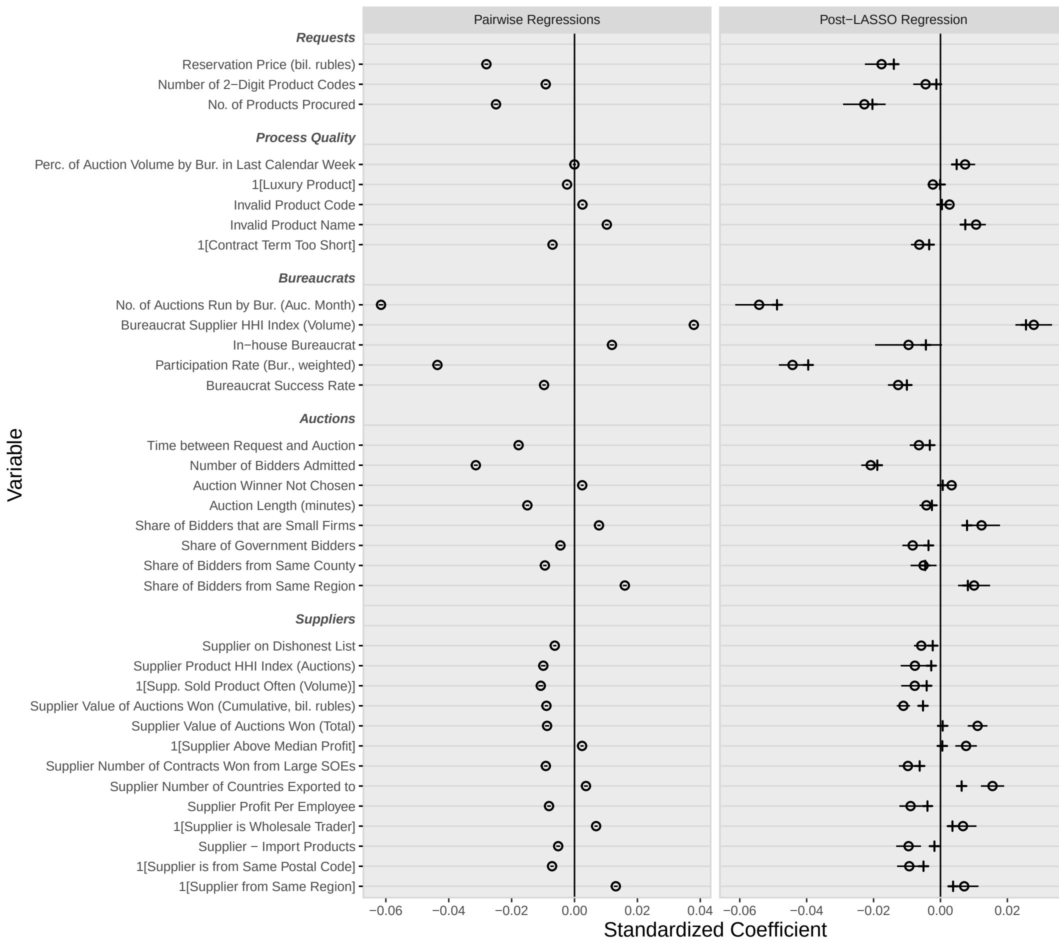


Figure 4. Correlates of Bureaucrat Effectiveness (Price)

Note: The figure shows the results of regressions of estimated bureaucrat effects $\hat{\alpha}_b$ from estimation of equation (3) for prices: $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ on observable characteristics of the purchase procedure followed. As described in section IV.E, since our data contain a large number of observables (see Table F.1), we use a LASSO procedure to select 30 predictor variables and regress each purchase's covariance-shrunk bureaucrat effect on these variables, the purchase's organization effect, and the controls in (3). The left panels show regression coefficients (in circles) and confidence intervals from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients (in circles) and confidence from the multivariate regression of the effects on all of the selected variables as well as the LASSO coefficients (as crosses). To facilitate comparison of effect sizes across variables, all variables are standardized to have unit standard deviation.

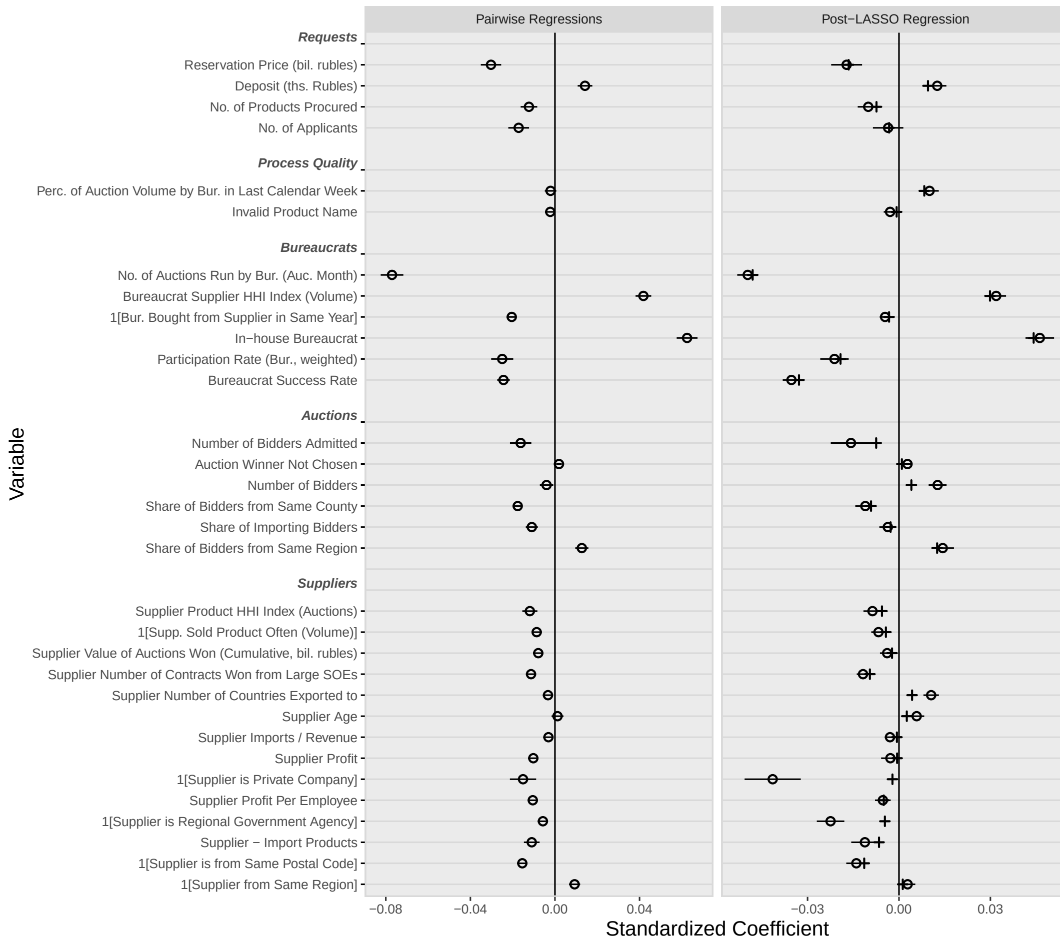


FIGURE 5. CORRELATES OF BUREAUCRAT EFFECTIVENESS (QUALITY)

Note: The figure shows the results of regressions of estimated bureaucrat effects $\hat{\alpha}_b$ from estimation of equation (3) for spending quality as discussed in Section IV.D: $q_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ on observable characteristics of the purchase procedure followed. As described in section IV.E, we use a LASSO procedure to select 30 predictor variables and regress each purchase's covariance-shrunk bureaucrat effect on these variables, the purchase's organization effect, and the controls in (3). The left panels show regression coefficients (in circles) and confidence intervals from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients (in circles) and confidence from the multivariate regression of the effects on all of the selected variables as well as the LASSO coefficients (as crosses). To facilitate comparison of effect sizes across variables, all variables are standardized to have unit standard deviation.

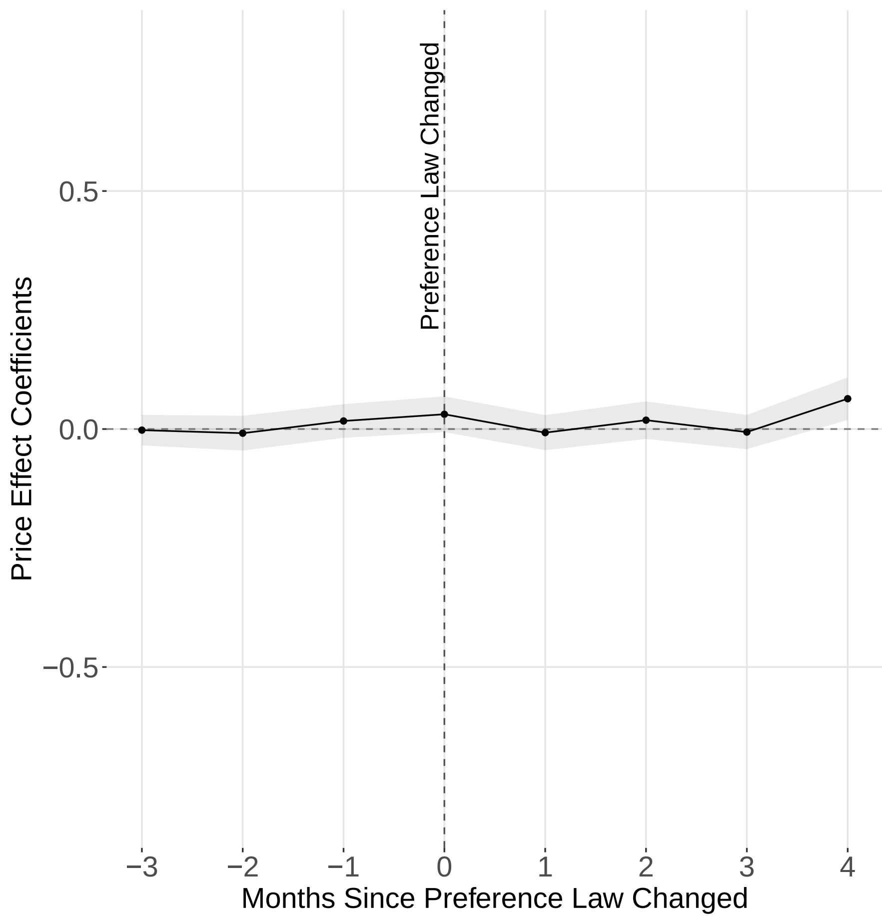
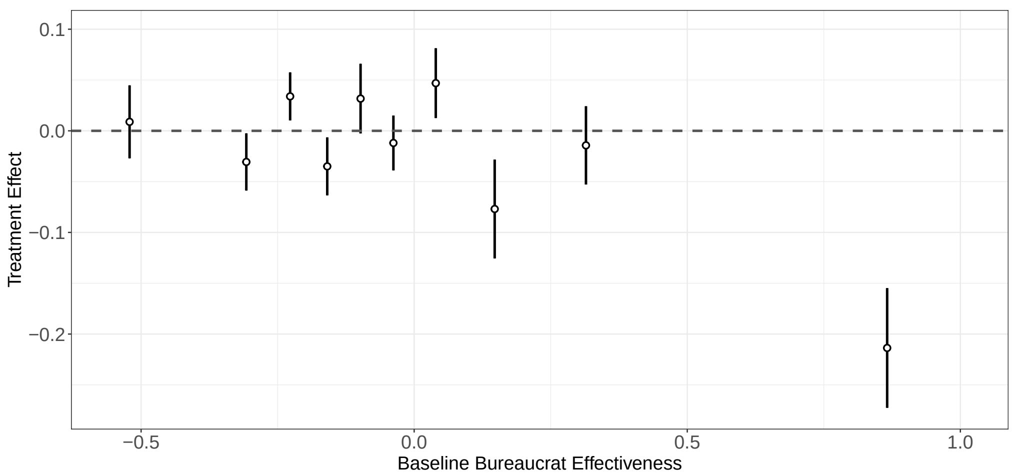


FIGURE 6. EVENT STUDY OF EFFECT OF BID PREFERENCES ON AVERAGE PRICES

Note: The figure shows the results of an event study analysis of the impact of the preferences policy on prices. Following Cengiz et al. (2019), we stack all the events (the preference list being published) and focus on a window starting three months before and ending four months after each year's preference list is published. We estimate equation (6): $p_{igt} = \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \sum_{s=-3}^{4} \delta_s \text{Preferenced}_{gt} \times \mathbf{1}\{t - \text{ListMonth}_t = s\} + \varepsilon_{igt}$ where p_{igt} is the log price paid in transaction i for good g in month t; $\mathbf{X}_{igt}$ are the same controls we use in Section IV, but for clarity we separate out the good and month fixed effects, μ_g and λ_t; Preferenced_{gt} is a dummy indicating that g is on the preferences list in the year month t falls within, ListMonth_t is the month closest to month t in which a preference list is published; and ε_{igt} is an error term we allow to be clustered by month and good. The figure shows the estimated δ_s coefficients and their 95% confidence intervals.

Panel A: Difference in Differences by Bureaucrat Effectiveness Decile



Panel B: Event Study by Bureaucrat Effectiveness

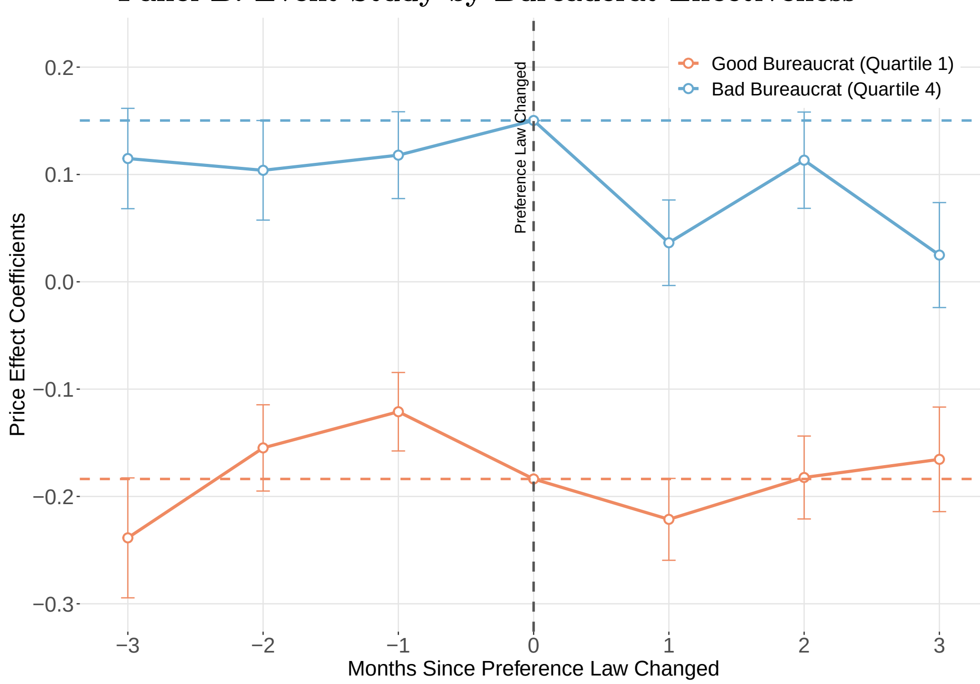


FIGURE 7. HETEROGENEITY OF BID PREFERENCES' EFFECT BY BUREAUCRAT EFFECTIVENESS

Note: The figure shows how the impacts of the introduction of bid preferences varies by the effectiveness of the implementing bureaucrat. Panel A shows estimates from implementing the triple difference model (8) to estimate separate effects for each decile of bureaucrat effectiveness: $y_{igt} = \sum_{k=1}^{10} \left\{ D_{kj} + D_{kb} \times (\rho_k \text{Preferenced}_{gt} + \eta_k \text{PolicyActive}_t + \pi_k \text{Preferenced}_{gt} \times \text{PolicyActive}_t) \right\} + \mathbf{X}_{igt}\beta + \mu_g + \mu_t + \varepsilon_{igt}$ where D_{kj} and D_{kb} are indicators for organization j and bureaucrat b belonging to decile k of their respective distributions of effectiveness. The horizontal axis plots the average effectiveness within the relevant decile, while the vertical axis plots the estimated treatment effects π_k with their 95% confidence intervals. Panel B extends the event study (6) shown in figure 6 (see notes to figure 6 for details) to estimate separate effects for the top and bottom quartile of bureaucrats. Rather than normalizing the reference month (the month before the preference list is published) to zero, we normalized it to the baseline performance in each group to better highlight how different their performance was before the preferences were introduced, and how their performance converges as a result of the preferences.

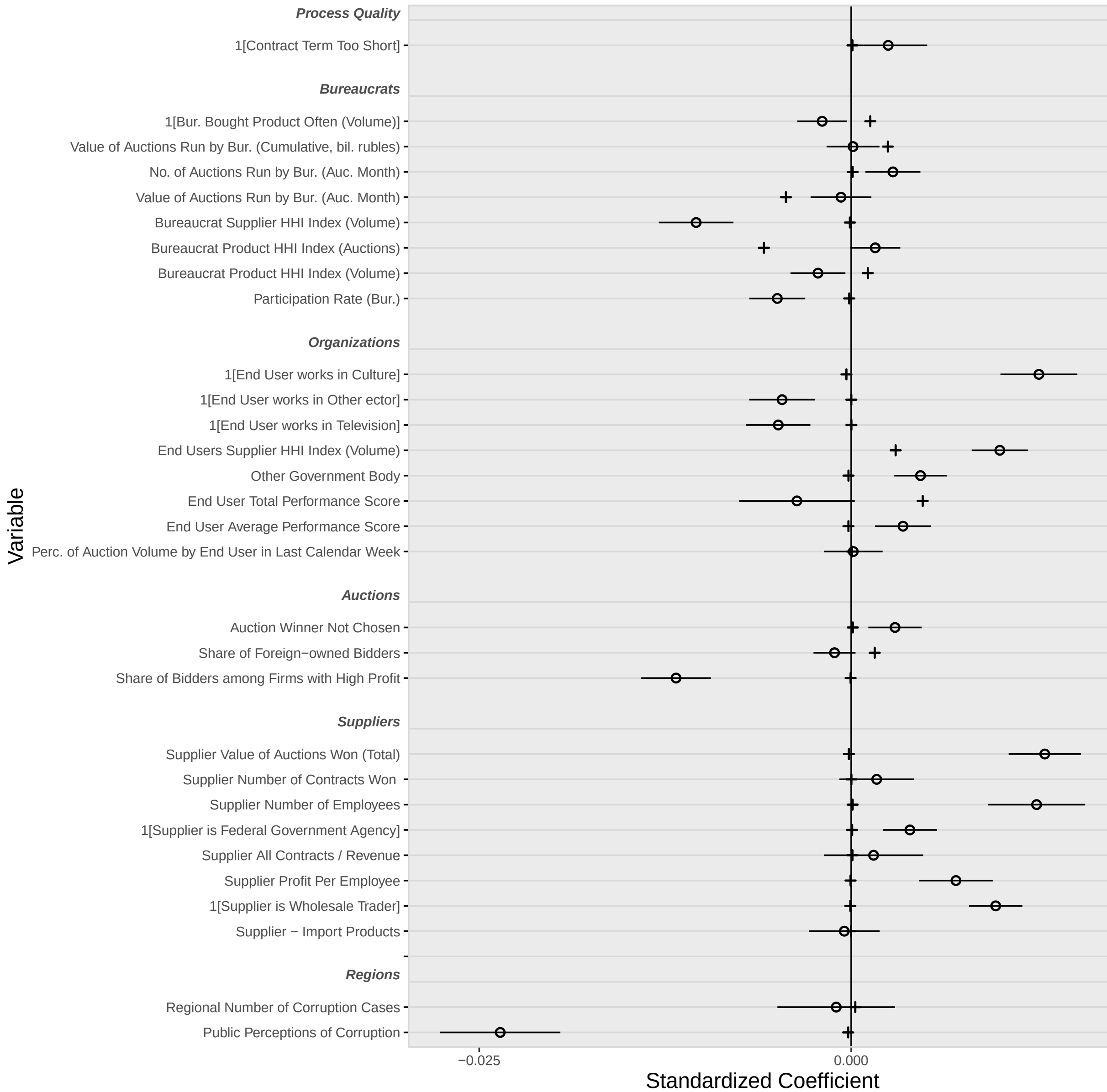


FIGURE 8. PREDICTORS OF HETEROGENEITY OF EFFECT OF BID PREFERENCES FOR DOMESTIC PRODUCERS

Note: The figure shows the results of estimating our triple-differences specification for heterogeneity of the effect of bid preferences (9): $y_{igt} = \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \mathbf{Z}_{igt}\boldsymbol{\theta} + \text{Preferenced}_{gt} \times \mathbf{Z}_{igt}\boldsymbol{\gamma} + \text{PolicyActive}_t \times \mathbf{Z}_{igt}\boldsymbol{\eta} + \delta\text{Preferenced}_{gt} \times \text{PolicyActive}_t + \text{Preferenced}_{gt} \times \text{PolicyActive}_t \times \mathbf{Z}_{igt}\boldsymbol{\pi} + \varepsilon_{igt}$. Since our data contain a large number of these (see Table F.1), the vector $\mathbf{Z}_{igt}$ is chosen by the same regularization procedure used in figure 4 and described in Sub-section IV.E to return 30 non-zero coefficients. The coefficients from the LASSO are shown as crosses, while the circles show the coefficients and 95% confidence intervals of a multivariate regression including the 30 observables.

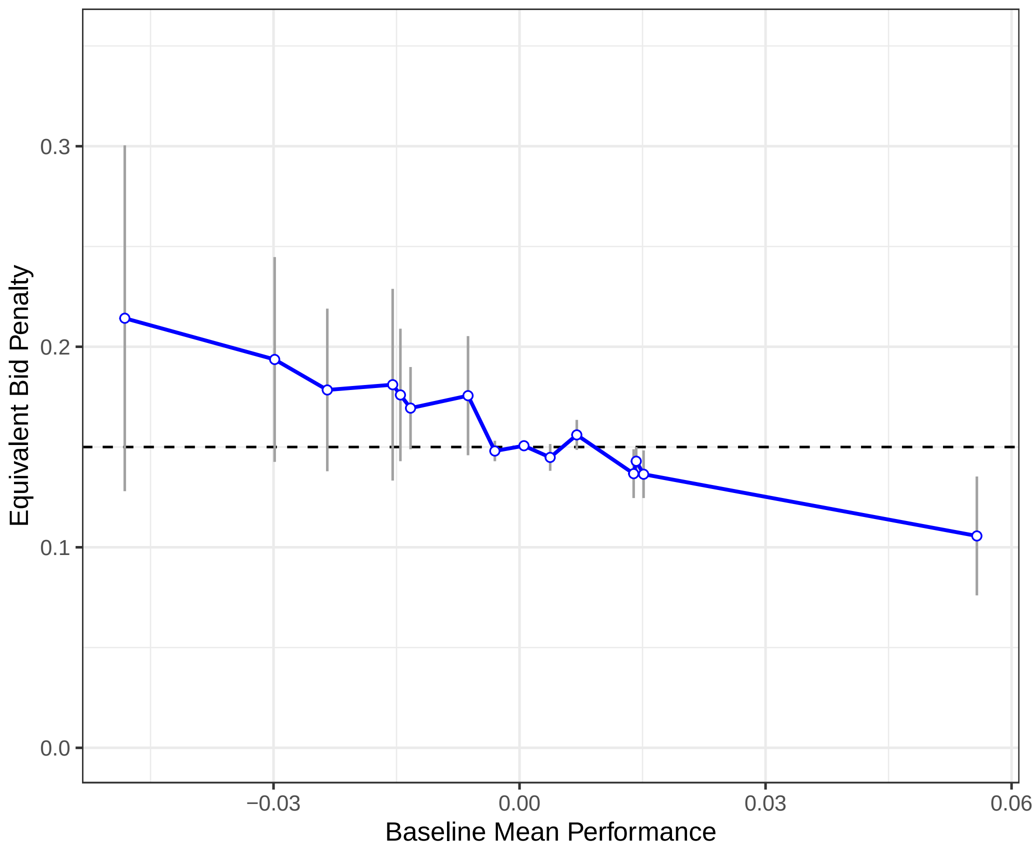


Figure 9. Tailoring Bid Preferences to Bureaucratic Capacity

Note: The figure shows estimates of the preference policy that attains the same impact in subsamples of bureaucrats as the 15 percent policy achieves in the overall sample. We combine our estimates in figure 7 of the effect of the 15 percent preference in each decile of the overall effectiveness distribution with the distributions of bureaucratic effectiveness in a range of subgroups of our data. The subgroups considered are, from left to right: sport department; culture department; regional government; housing department; non in-house bureaucrats; bureaucrats with high auction volume; buyers far from their regional capital; buyers near their regional capital; bureaucrats with low auction volume; in-house bureaucrats; municipal government; education department; internal affairs department; other departments; and federal government. We assume that for each decile k of effectiveness, log prices are locally linear in the preference rate, with slope $TE_k/0.15$, where TE_k is the treatment effect for decile k estimated using equation (8) shown in Figure 7A. For any subgroup g with a distribution of bureaucrats $w_{kg}, j = 1, \ldots, 10$ across the deciles of effectiveness, we can find the preference rate $1 - \gamma_g^*$ that would achieve the same impact in that subgroup as the 15 percent rate achieves in the overall sample as follows. For each subgroup, our estimates in figure 7A imply a treatment effect of a 15 percent bid penalty of $TE_g = \sum_{k=1}^{10} w_{jg} TE_k$, and our constant elasticity assumption implies that the equivalent policy solves (10) yielding $1 - \gamma_g^* = 0.15\overline{TE}/TE_g$ where $\overline{TE} = \sum_{k=1}^{10} TE_k$ is the treatment effect in the overall sample. These are shown on the vertical axis of the figure along with their 95% confidence intervals

Table 1—Summary Statistics

	All Products			Pharmaceuticals Subsample		
	(1) No Preferences Full Sample	(2) No Preferences Analysis Sample	(3) Analysis Sample With Preferences	(4) No Preferences Full Sample	(5) No Preferences Analysis Sample	(6) Analysis Sample With Preferences
(1) # of Bureaucrats	115,854	37,722	37,722	5,560	2,473	2,473
(2) # of Organizations	88,306	44,560	44,560	3,662	1,866	1,866
(3) # of Connected Sets	26,234	616	616	0	129	129
(4) # of Bureaucrats with >1 Org.	14,090	11,063	11,320	965	926	1,095
(5) # of Organizations with >1 Bur.	54,575	37,306	37,536	2,076	1,449	1,596
(6) Mean # of Bureaucrats per Org.	3.96	5.59	6.02	3.1	4.32	6.3
(7) Mean # of Organizations per Bur.	3.02	6.6	7.12	2.04	3.26	4.75
(8) # of Federal Organizations	12,889	1,583	1,583	496	26	26
(9) # of Regional Organizations	25,162	15,530	15,530	2,786	1,599	1,599
(10) # of Municipal Organizations	50,220	27,440	27,440	380	241	241
(11) # of Health Organizations	10,167	7,231	7,231	3,172	1,705	1,705
(12) # of Education Organizations	42,045	25,271	25,271	109	61	61
(13) # of Internal Affairs Organizations	3,126	668	668	105	3	3
(14) # of Agr/Environ Organizations	1,032	255	255	26	1	1
(15) # of Other Organizations	31,936	11,135	11,135	250	96	96
(16) # of Goods	16,373	14,875	15,649	4,220	3,861	4,351
(17) Mean # of Goods Per Bur.	35	72.5	93.2	31.6	42.5	82.3
(18) # of Regions	86	86	86	85	79	79
(19) Mean # of Regions per Bur.	1	1	1	1	1	1
(20) # of Auction Requests	1,733,433	1,199,363	1,871,717	62,755	42,875	114,808
(21) Mean # of Requests per Bur.	15	31.8	49.6	11.3	17.3	46.4
(22) Mean # of Applicants	3.01	3.04	2.94	2.57	2.65	2.7
(23) Mean # of Bidders	2.06	1	2.07	1.94	1.98	2
(24) Mean Reservation Price	0.29	0.291	0.291	0.303	0.303	0.302
(25) Quantity Mean	1,131	1,053	1,124	1,201	1,719	975
Median	20	25	27	40	45	50
SD	80,563	90,917	174,951	136,260	172,144	108,598
(26) Total Price Mean (bil. USD)	93.3	80.1	81.2	128	91.1	101
Median	4.67	4.32	4.74	6.23	6.7	7.06
SD	578	493	482	5,745	493	525
(27) Unit Price Mean (bil. USD)	72.1	61.3	55.6	20.2	25.4	28.8
Median	0.21	0.167	0.18	0.175	0.18	0.18
SD	21,248	23,015	19,168	226	265	281
(28) Mean # of Contract Renegotiations (log)	0.126	0.121	0.133	0.15	0.142	0.178
(29) Mean Size of Cost Over-run	-0.001	-0.002	-0.002	-0.003	-0.003	-0.003
(30) Mean Length of Delay in Days (log)	0.061	0.064	0.057	0.077	0.076	0.069
(31) Mean 1[End User Complained about Contract]	0.001	0.001	0.001	0.001	0	0.001
(32) Mean 1[Contract Cancelled]	0.012	0.012	0.009	0.016	0.016	0.01
(33) Mean 1[Product is of Substandard Quality]	0.005	0.005	0.009	0.075	0.058	0.041
(34) # of Observations	15,096,253	11,339,187	16,348,331	290,483	181,963	460,533
(35) Total Procurement Volume (bil. USD)	516	395	629	14.5	9.38	19.9

Note: The table reports summary statistics for six samples. The All Products columns show statistics for purchases of all off-the-shelf goods, while the Pharmaceuticals Subsample columns restrict attention to purchases of medicines. Full Sample denotes all unpreferenced auctions. Analysis Sample denotes all unpreferenced auctions in connected sets that fulfill the restrictions discussed in section IV.B: singleton bureaucrat-organization, bureaucrat-good, organization-good pairs, and levels of our control fixed effects are removed; each procurer (bureaucrats and organizations) implements a minimum of five purchases; and connected sets have at least three bureaucrats and organizations. With Bid Preferences denotes all preferenced auctions that fulfill the same three restrictions. Organizations working in Education include schools, universities, pre-schools, and youth organizations. Organizations working in Internal Affairs include police, emergency services, local administration, taxes, and transportation. Organizations working in Agriculture or the Environment include environmental protection funds, agricultural departments and nature promotion agencies. The Other category includes funds, monitoring agencies, and land cadasters, among many others. All sums are measured in billions of US dollars at an exchange rate of 43 rubles to 1 US dollar.

Table 2—Share of Variation in Policy Performance Explained by Bureaucrats and Organizations

	Fixed Effects (1)	(s.e.) (2)	Split Sample (3)	(s.e.) (4)	Shrinkage (5)	Covariance Shrinkage (6)
(1) s.d. of Bureaucrat Effects (across burs)	1.197	(0.030)	1.259	(0.0301)	0.824	0.423
(2) s.d. of Organization Effects (across orgs)	1.130	(0.0412)	1.184	(0.049)	0.785	0.368
(3) s.d. of Bureaucrat Effects (across items)	0.788	(0.0316)	0.834	(0.0443)	0.595	0.263
(4) s.d. of Organization Effects (across items)	0.927	(0.0464)	0.976	(0.0577)	0.709	0.338
(5) Bur-Org Effect Correlation (across items)	-0.720	(0.0173)	-0.561	(0.0375)	-0.663	0.311
(6) s.d. of Bur + Org Effects Within CS (across items)	0.655	(0.0185)	0.661	(0.0198)	0.545	0.489
(7) s.d. of log unit price	2.188		2.188		2.188	2.188
(8) s.d. of log unit price | good, month	1.280		1.280		1.280	1.280
(9) Adjusted R-squared	0.963		0.963		0.963	0.963
(10) Number of Bureaucrats	37,722		37,722		37,722	37,722
(11) Number of Organizations	44,560		44,560		44,560	44,560
(12) Number of Bureaucrat-Organization Pairs	248,898		248,898		248,898	248,898
(13) Number of Connected Sets	616		616		616	616
(14) Number of Observations	11,339,187		11,339,187		11,339,187	11,339,187

Note: The table shows the components of the variance due to bureaucrats, organizations, and connected sets estimated by implementing the variance decomposition in equation (4). The sample used is the All Products-Analysis Sample summarized in Table 1. Rows 1 & 2 show the s.d. of the bureaucrat, organization and connected set effects. Rows 3–6 show the components of the variance of prices across purchases, effectively weighting the estimates in rows 1–3 by the number of purchases they conduct. Column 1 uses the fixed effect estimates from equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. Each observation is an item procured by an organization j and a bureaucrat indexed by $b(i,j)$. Column 3 shows estimates from randomly splitting the sample in half, stratifying by bureaucrat-organization pair and calculating the covariance across the two noisy estimates. Columns 2 and 4 show standard errors of the estimates in columns 1 and 3, respectively, estimated by bootstrapping 100 times. Column 5 uses the bootstraps to estimate the sampling error in each bureaucrat effect s_b^2 and each organization effect s_j^2, and the signal variances of the bureaucrat and organization effects (σ_α^2 and σ_ψ^2 respectively). The minimum-mean-squared error predictor for each bureaucrat effect is then $[\hat{\sigma}_\alpha^2/(\hat{\sigma}_\alpha^2 + s_b^2)] \cdot \hat{\alpha}_b$, where $\hat{\alpha}_b$ is the bureaucrat's fixed effect from the decomposition in Column 1, and analogously for the organization effects. Column 6 shows our preferred estimates, which form predictions of the bureaucrat and organization effects that minimize the expected sum of the mean-squared errors of the predictions and take into account the covariance of the estimation errors, estimated from the bootstrapped estimates. Formally, the covariance shrinkage predictors solve $\min_{\boldsymbol{\Lambda}} \mathbb{E}\left[\left(\boldsymbol{\theta} - \boldsymbol{\Lambda}\hat{\boldsymbol{\theta}}\right)'\left(\boldsymbol{\theta} - \boldsymbol{\Lambda}\hat{\boldsymbol{\theta}}\right)\right]$ where $\hat{\boldsymbol{\theta}}$ is the vector of estimated bureaucrat and organization fixed effects. All methods are described fully in Section IV.B.

Table 3—Robustness to Restricting to Pharmaceuticals Subsample with Barcode Information

	Fixed Effects (1)	(s.e.) (2)	Split Sample (3)	(s.e.) (4)	Shrinkage (5)	Covariance Shrinkage (6)
(1) s.d. of Bureaucrat Effects (across burs)	0.266	(0.0116)	0.244	(0.0152)	0.124	0.0803
(2) s.d. of Organization Effects (across orgs)	0.207	(0.00752)	0.210	(0.00792)	0.0883	0.0575
(3) s.d. of Bureaucrat Effects (across items)	0.184	(0.0293)	0.191	(0.0317)	0.110	0.0665
(4) s.d. of Organization Effects (across items)	0.197	(0.0359)	0.206	(0.0376)	0.106	0.0544
(5) Bur-Org Effect Correlation (across items)	-0.544	(0.0854)	-0.304	(0.0698)	-0.276	-0.0304
(6) s.d. of Bur + Org Effects Within CS (across items)	0.183	(0.00625)	0.183	(0.00687)	0.130	0.0846
(7) s.d. of log unit price	1.914		1.914		1.914	1.914
(8) s.d. of log unit price | good, month	0.430		0.430		0.430	0.430
(9) Adjusted R-squared	0.996		0.996		0.996	0.996
(10) Number of Bureaucrats	2,473		2,473		2,473	2,473
(11) Number of Organizations	1,866		1,866		1,866	1,866
(12) Number of Bureaucrat-Organization Pairs	8,067		8,067		8,067	8,067
(13) Number of Connected Sets	129		129		129	129
(14) Number of Observations	181,963		181,963		181,963	181,963

Note: The table shows the components of the variance due to bureaucrats, organizations, and connected sets estimated by implementing the variance decomposition in equation (4). The sample used is the Pharmaceuticals-Analysis Sample summarized in Table 1. The table is constructed analogously to table 2 (whose notes contain further details). All methods are described fully in Section IV.B.

Table 4—Spending Quality Variance Decomposition

	Fixed Effects (1)	(s.e.) (2)	Split Sample (3)	(s.e.) (4)	Shrinkage (5)	Covariance Shrinkage (6)
(1) s.d. of Bureaucrat Effects (across burs)	0.378	(0.0257)	0.423	(0.0276)	0.186	0.0995
(2) s.d. of Organization Effects (across orgs)	0.380	(0.0409)	0.424	(0.0459)	0.193	0.0884
(3) s.d. of Bureaucrat Effects (across items)	0.338	(0.048)	0.367	(0.0509)	0.185	0.0849
(4) s.d. of Organization Effects (across items)	0.373	(0.0496)	0.401	(0.0547)	0.209	0.0856
(5) Bur-Org Effect Correlation (across items)	-0.809	(0.0287)	-0.607	(0.0788)	-0.699	0.336
(6) s.d. of Bur + Org Effects Within CS (across items)	0.222	(0.0168)	0.226	(0.0143)	0.154	0.139
(7) s.d. of quality index	0.592		0.592		0.592	0.592
(8) s.d. of quality index | good, month	0.571		0.571		0.571	0.571
(9) Adjusted R-squared	0.946		0.946		0.946	0.946
(10) Number of Bureaucrats	37,722		37,722		37,722	37,722
(11) Number of Organizations	44,560		44,560		44,560	44,560
(12) Number of Bureaucrat-Organization Pairs	248,898		248,898		248,898	248,898
(13) Number of Connected Sets	616		616		616	616
(14) Number of Observations	11,339,187		11,339,187		11,339,187	11,339,187

Note: Notes: The table shows the components of the variance due to bureaucrats, organizations, and connected sets estimated by implementing the variance decomposition in equation (4) but with spending quality as the outcome, as discussed in section IV.D. The sample used is the All Products-Analysis Sample summarized in Table 1. Rows 1 & 2 show the s.d. of the bureaucrat, organization and connected set effects. Rows 3–6 show the components of the variance of prices across purchases, effectively weighting the estimates in rows 1 & 2 by the number of purchases they conduct. Column 1 uses the fixed effect estimates from equation (3): $q_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. Each observation is an item procured by an organization j and a bureaucrat indexed by $b(i,j)$. Column 3 shows estimates from randomly splitting the sample in half, stratifying by bureaucrat-organization pair and calculating the covariance across the two noisy estimates. Columns 2 and 4 show standard errors of the estimates in columns 1 and 3, respectively, estimated by bootstrapping 100 times. Column 5 uses the bootstraps to estimate the sampling error in each bureaucrat effect s_b^2 and each organization effect s_j^2, and the signal variances of the bureaucrat and organization effects (σ_α^2 and σ_ψ^2 respectively). The minimum-mean-squared error predictor for each bureaucrat effect is then $[\hat{\sigma}_\alpha^2/(\hat{\sigma}_\alpha^2 + s_b^2)] \cdot \hat{\alpha}_b$, where $\hat{\alpha}_b$ is the bureaucrat's fixed effect from the decomposition in Column 1, and analogously for the organization effects. Column 6 shows our preferred estimates, which form predictions of the bureaucrat and organization effects that minimize the expected sum of the mean-squared errors of the predictions and take into account the covariance of the estimation errors, estimated from the bootstrapped estimates. Formally, the covariance shrinkage predictors solve $\min_{\boldsymbol{\Lambda}} \mathbb{E}\left[\left(\boldsymbol{\theta} - \boldsymbol{\Lambda}\hat{\boldsymbol{\theta}}\right)'\left(\boldsymbol{\theta} - \boldsymbol{\Lambda}\hat{\boldsymbol{\theta}}\right)\right]$ where $\hat{\boldsymbol{\theta}}$ is the vector of estimated bureaucrat and organization fixed effects. All methods are described fully in Section IV.B.

Table 5—Bid Preferences Increase Domestic Winners With Limited Impact on Prices or Participation

	All Products			Pharmaceuticals			
	Log Price (1)	Num. Bidders (2)	Quality (3)	Log Price (4)	Num. Bidders (5)	Quality (6)	Domestic Winner (7)
log Standardized Quantity	-0.308***	0.043***	-0.001**	-0.030***	0.012***	-0.012***	0.004***
	(0.015)	(0.002)	(0.0006)	(0.002)	(0.002)	(0.0006)	(0.0008)
Preferenced * Policy Active	-0.004	-0.041***	-0.022***	-0.025**	-0.024	0.013*	0.042***
	(0.010)	(0.011)	(0.005)	(0.011)	(0.020)	(0.007)	(0.005)
R^2	0.653	0.266	0.226	0.948	0.271	0.257	0.736
Observations	16,348,331	16,348,331	16,348,331	460,533	460,533	460,533	460,533
Outcome Mean	5.557	2.065	0.075	6.279	1.942	0.178	0.385
Constituent Terms	✓	✓	✓	✓	✓	✓	✓
Good fixed effects	✓	✓	✓	✓	✓	✓	✓
Month fixed effects	✓	✓	✓	✓	✓	✓	✓
$Year * Product * Size * Region$ fixed effects	✓	✓	✓	✓	✓	✓	✓

Note: This table estimates the Intent to Treat (ITT) of the bid preference policy from equation (5): $y_{igt} = \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \delta \text{Preferenced}_{gt} \times \text{PolicyActive}_t + \varepsilon_{igt}$. The sample used is summarized in columns (3) and (6) of Table 1. In the All Products sample an item has $\text{Preferenced}_{gt} = 1$ if the type of good appears on the list of goods covesred by the preferences policy for that year. In the Pharmaceuticals sample, $\text{Preferenced}_{gt} = 1$ if the drug purchased is made both in Russia and abroad. $\text{PolicyActive}_t = 1$ during the part of the relevant year that the preferences policy was in effect. Standard errors are clustered by month and good.

Table 6—Bid Preferences are More Effective When Implemented by Less Effective Bureaucrats

	All Products			Pharmaceuticals			
	Log Price (1)	Num. Bidders (2)	Quality (3)	Log Price (4)	Num. Bidders (5)	Quality (6)	Domestic Winner (7)
log Standardized Quantity	-0.309***	0.042***	-0.0006	-0.027***	0.008***	-0.007***	0.004***
	(0.015)	(0.002)	(0.0006)	(0.002)	(0.002)	(0.0006)	(0.0008)
Bureaucrat FE ∗ Preferenced ∗ Policy Active	-0.090***	0.081***	-0.051	-0.466***	1.12***	0.064	0.211***
	(0.018)	(0.018)	(0.037)	(0.090)	(0.201)	(0.067)	(0.037)
Organization FE ∗ Preferenced ∗ Policy Active	0.004	0.005	0.052	0.084	0.450**	0.004	-0.145***
	(0.017)	(0.013)	(0.036)	(0.103)	(0.193)	(0.090)	(0.036)
R^2	0.658	0.271	0.243	0.950	0.288	0.320	0.736
Observations	16,348,331	16,348,331	16,348,331	460,533	460,533	460,533	460,533
Outcome Mean	5.557	2.065	0.075	6.279	1.942	0.178	0.385
Constituent Terms	✓	✓	✓	✓	✓	✓	✓
Good fixed effects	✓	✓	✓	✓	✓	✓	✓
$Year * Product * Size * Region$ fixed effects	✓	✓	✓	✓	✓	✓	✓
Month fixed effects	✓	✓	✓	✓	✓	✓	✓
ConnectedSet fixed effects	✓	✓	✓	✓	✓	✓	✓

Note: This table estimates the triple-difference from equation (7): $y_{igt} = \mathbf{X}_{igt}\beta + \mu_g + \lambda_t + \theta_b\hat{\alpha}_b + \theta_j\hat{\psi}_j + \delta\text{Preferenced}_{gt} \times \text{PolicyActive}_t + \rho_b\text{Preferenced}_{gt}\hat{\alpha}_b + \rho_j\text{Preferenced}_{gt}\hat{\psi}_j + \eta_b\text{PolicyActive}_t\hat{\alpha}_b + \eta_j\text{PolicyActive}_t\hat{\psi}_j + \pi_b\text{Preferenced}_{gt} \times \text{PolicyActive}_t\hat{\alpha}_b + \pi_j\text{Preferenced}_{gt} \times \text{PolicyActive}_t\hat{\psi}_j + \varepsilon_{igt}$. The sample used is summarized in columns (3) and (6) of Table 1. In the All Products sample an item has $\text{Preferenced}_{gt} = 1$ if the type of good appears on the list of goods covered by preferences that year. In the Pharmaceuticals sample, $\text{Preferenced}_{gt} = 1$ if the drug purchased is made both in Russia and abroad. $\text{PolicyActive}_t = 1$ during the part of the year that the preferences policy was in effect. Bureaucrat and Organization FEs are the covariance-shrunk bureaucrat and organization effects estimated in section IV. Standard errors are clustered by month and good.