Online Appendix: *Individuals and Organizations as Sources of State Effectiveness*, Michael Carlos Best, Jonas Hjort, & David Szakonyi

Table of Contents

A: Details on Text Analysis

This appendix provides some of the details of the procedure we use to categorize procurement purchases into groups of homogeneous products. We proceed in three steps. First, we transform the raw product descriptions in our data into vectors of word tokens to be used as input data in the subsequent steps. Second, we develop a transfer learning procedure to use product descriptions and their corresponding Harmonized System product codes in data on the universe of Russian imports and exports to train a classification algorithm to assign product codes to product descriptions. We then apply this algorithm to the product descriptions in our procurement data. Third, for product descriptions that are not successfully classified in the second step, either because the goods are non-traded, or because the product description is insufficiently specific, we develop a clustering algorithm to group product descriptions into clusters of similar descriptions.

Once our data is grouped into products, we create our main outcome of interest–unit prices—in three steps. First, we standardize all units to be in SI units (e.g. convert all lengths to meters). Second, for each good, we keep only the most frequent standardized units i.e. if a good is usually purchased by weight and sometimes by volume, we keep only purchases by weight. Third, we drop the top and bottom 5% of the unit prices for each good since in some cases the number of units purchased is off by an order of magnitude spuriously creating very large or very small unit prices due to measurement error in the quantity purchased.

A1 Preparing Text Data

The first step of our procedure 'tokenizes' the sentences that we will use as inputs for the rest of the procedure. We use two datasets of product descriptions. First, we use the universe of customs declarations on imports and exports to & from Russia in 2011–2013. Second, we use the product descriptions in our procurement data described in Subsection II.A. Each product description is parsed in the following way, using the Russian libraries for Python's Natural Language Toolkit[73]

1) Stop words are removed that are not core to the meaning of the sentence, such as "the", "and", and "a".
2) The remaining words are lemmatized, converting all cases of the same word into the same 'lemma' or stem. For example, 'potatoes' become 'potato'.
3) Lemmas two letters or shorter are removed.

We refer to the result as the *tokenized* sentence. For example the product description "NV-Print Cartridge for the Canon LBP 2010B Printer" would be broken into the following tokens: [cartridge, NV-Print, printer, Canon, LBP, 3010B].

[73]Documentation on the Natural Language Toolkit (NLTK) can be found at http://www.nltk.org/

[74] Similarly, the product description "sodium bicarbonate - solution for infusion 5%,200ml" would result in the following tokens: [sodium, bicarbonate, solution, infusion, 5%, 200ml].[75]

A2 Classification

In the second step of our procedure we train a classification algorithm to label each of the sentences in the customs data with one of the H_C labels in the set of labels in the customs dataset, $\mathcal{H}_C$. To prepare our input data, each of the N_C tokenized sentences $\mathbf{t}_i$ in the customs dataset is transformed into a vector of token indicators and indicators for each possible bi-gram (word-pair), denoted by $\mathbf{x}_i \in \mathcal{X}_C$.[76] Each sentence also has a corresponding good classification $g_i \in \mathcal{G}_C$, so we can represent our customs data as the pair $\{\mathbf{X}_C, \mathbf{g}_C\}$ and we seek to find a classifier $\hat{g}_C(\mathbf{x}) : \mathcal{X}_C \to \mathcal{H}_C$ that assigns every text vector $\mathbf{x}$ to a product code.

As is common in the literature, rather than solving this multiclass classification problem in a single step, we pursue a "one-versus-all" approach and reduce the problem of choosing among G possible good classifications to G_C binary choices between a single good and all other goods, and then combine them (Rifkin and Klautau, 2004). We do this separately for each 2-digit product category. Each of the G_C binary classification algorithms generates a prediction $p_g(\mathbf{x}_i)$, for whether sentence i should be classified as good g. We then classify each sentence as the good with the highest predicted value:

$$\hat{g}_C(\mathbf{x}_i) = \arg\max_{g \in \mathcal{G}_C} p_g(\mathbf{x}_i) \tag{A.1}$$

Each binary classifier is a logistic regression solving

$$\min_{\mathbf{w}_g, a_g} \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{1}{\ln 2} \ln\left(1 + \mathrm{e}^{-y_{gi} \cdot (\mathbf{w}_g \cdot \mathbf{x}_i + a_g)}\right) \tag{A.2}$$

where

$$y_{gi} = \begin{cases} 1 & \text{if } g_i = g \\ -1 & \text{otherwise} \end{cases}$$

The minimands $\hat{\mathbf{w}}_g$ and $\hat{a}_g$ are then used to compute $p_g(\mathbf{x}_i) = \hat{\mathbf{w}}_g \cdot \mathbf{x}_i + \hat{a}_g$ with which the final classification is formed using equation (A.1). We implement this procedure using the Vowpal Wabbit library for Python.[77] This simple procedure is

[74] The original Russian text reads as "картридж NV-Print для принтера Canon LBP 3010B" with the following set of Russian tokens: [картридж, NV-Print, принтер, Canon, LBP, 3010B].

[75] The original Russian text reads as "натрия гидрокарбонат - раствор для инфузий 5%,200мл" with the set of Russian tokens as: [натрия, гидрокарбонат, раствор, инфузия, 5%, 200мл].

[76] The customs entry "Electric Table Lamps Made of Glass" is transformed into the set of tokens: [electric, table, lamp, glass]. The original Russian reads as "лампы электрические настольные из стекла" and the tokens as: [электрический, настольный, ламп, стекло].

[77] See http://hunch.net/~vw/.

remarkably effective; when trained on a randomly selected half of the customs data and then implemented on the reamining data for validation, the classifications are correct 95% of the time. Given this high success rate without regularization, we decided not to try and impose a regularization penalty to improve out of sample fit. We also experimented with two additional types of classifiers. First, we trained a linear support vector machine with a hinge loss function.[78] That is, a classifier that solves

$$\min_{\mathbf{w}_g, a_g} \frac{1}{N_C} \sum_{i=1}^{N_C} \max\left\{0, 1 - y_{gi} \cdot (\mathbf{w}_g \cdot \mathbf{x}_i + a_g)\right\} \tag{A.3}$$

Second, we trained a set of hierarchical classifiers exploiting the hierarchical structure of the HS product classification. Each classifier is a sequence of sub-classifiers. The first sub-classifier predicts which 4-digit HS code corresponds to the text. Then, within each 4-digit code, the next classifier predicts the corresponding 6-digit code, etc, until the last classifier that predicts the full 10-digit code within each 8-digit category. Our main analysis of section IV.C presented in figure 1 and table 2 is repeated using these alternative classifiers in figure D.1 panels C and D and in table E.4. As they show, the results are robust to these alternative classification methods.

Having trained the algorithm on the customs dataset, we now want to apply it to the procurement dataset wherever possible. This is known as transfer learning (see, for example Torrey and Shavlik (2009)). Following the terminology of Pang and Yang (2010), our algorithm $\hat{g}_C$ performs the task $\mathcal{T}_C = \{\mathcal{H}_C, g_C(\cdot)\}$ learning the function $g_C(\cdot)$ that maps from observed sentence data X to the set of possible customs labels $\mathcal{G}_C$. The algorithm was trained in the domain $\mathcal{D}_C = \{\mathcal{X}_C, F(X)\}$ where $F(\mathbf{X})$ is the probability distribution of $\mathbf{X}$. We now seek to transfer the algorithm to the domain of the procurement dataset, $\mathcal{D}_B = \{\mathcal{X}_B, F(\mathbf{X})\}$ so that it can perform the task $\mathcal{T}_B = \{\mathcal{H}_B, g_B(\cdot)\}$. Examples of the classification outcomes can be found in Tables A.1 (translated into English) and A.2 (in the original Russian). The three columns on the left present the tokens from the descriptions of goods in the procurement data, along with an identifying contract number and the federal law under which they were concluded. The columns on the right indicate the 10-digit HS code ('13926100000 - Office or school supplies made of plastics') that was assigned to all four of the goods using the machine learning algorithm. In addition, we present the tokenized customs entries that correspond to this 10 digit HS code.

The function to be learned and the set of possible words used are unlikely to differ between the two domains—A sentence that is used to describe a ball bearing in the customs data will also describe a ball bearing in the procurement data—so

[78]A description of the support vector loss function (hinge loss), which estimates the mode of the posterior class probabilities, can be found in Friedman, Hastie and Tibshirani (2013, 427)

Table A.1—Example Classification - English

Contract ID	Law	Product Description	HS10 Code	Example Import Entries
5070512	94FZ	folder, file, Erich, Krause, Standard, 3098, green	3926100000	product, office, made of, plastic
15548204	44FZ	cover, plastic, clear	3926100000	office, supply, made of, plastic, kids, school, age, quantity
16067065	44FZ	folder, plastic	3926100000	supply, office, cover, plastic, book
18267299	44FZ	folder, plastic, Brauberg	3926100000	collection, office, desk, individual, plastic, packaging, retail, sale

$\mathcal{X}_C = \mathcal{X}_B$, and $h_C(\cdot) = h_B(\cdot)$. The two key issues that we face are first, that the likelihoods that sentences are used are different in the two samples so that $F(\mathbf{X})_C \neq F(\mathbf{X})_B$. This could be because, for example, the ways that importers and exporters describe a given good differs from the way public procurement officials and their suppliers describe that same good. In particular, the procurement sentences are sometimes not as precise as those used in the trade data. The second issue is that the set of goods that appear in the customs data differs from the goods in the procurement data so that $\mathcal{H}_C \neq \mathcal{H}_B$. This comes about because non-traded goods will not appear in the customs data, but may still appear in the procurement data.

To deal with these issues, we identify the sentences in the procurement data that are unlikely to have been correctly classified by $\hat{h}_C$ and instead group them into goods using the clustering procedure described in section A.A3 below. We construct 2 measures of the likelihood that a sentence is correctly classified. First, the predicted value of the sentence's classification $\hat{g}_C(\mathbf{x}_i)$ as defined in (A.1). Second, the similarity between the sentence and the average sentence with the sentence's assigned classification in the *customs* data used to train the classifier.

To identify outlier sentences, we take the tokenized sentences that have been labeled as good g, $\mathbf{t}_g = \{\mathbf{t}_i : \hat{g}_C(\mathbf{x}_i) = g\}$ and transform them into vectors of indicators for the tokens $\mathbf{v}_{gi}$.[79] For each good, we then calculate the mean sentence vector in the customs data as $\mathbf{v}_g^C = \sum_{\mathbf{v}_{gi}, x_i \in \mathbf{X}^C} \mathbf{v}_{gi} / |\mathbf{t}_g|$. Then, to identify outlier sentences in the procurement data, we calculate each sentence's normalized cosine

[79] Note that these vectors differ from the inputs $\mathbf{x}_i$ to the classifier in two ways. First, they are specific to a certain good, and second, they omit bigrams of the tokens

Table A.2—Example Classification - Russian

Contract ID	Law	Product Description	HS10 Code	Example Import Entries
5070512	94FZ	Папка, файл, Erich, Krause, Standard, 3098, зелёная	3926100000	изделие, канцелярский, изготовленный, пластик
15548204	44FZ	Обложка, пластиковый, прозрачный	3926100000	канцелярский, принадлежность, изготовленный, пластик, дети, школьный, возрасть, количество
16067065	44FZ	Скоросшиватель, пластиковый	3926100000	принадлежность, канцелярский, закладка, пластиковый, книга
18267299	44FZ	Скоросшиватель, пластиковый, Brauberg	3926100000	набор, канцелярский, настольный, индивидуальный, пластмассовый, упаковка, розничный, продажа

similarity with the good's mean vector,

$$\theta_{gi} = \frac{\bar{s}_g - s\left(\mathbf{v}_{gi}, \mathbf{v}_g\right)}{\bar{s}_g} \tag{A.4}$$

where $s\left(\mathbf{v}_{gi}, \mathbf{v}_g\right) \equiv \cos\left(\mathbf{v}_{gi}, \mathbf{v}_g\right) = \frac{\mathbf{v}_{gi}\mathbf{v}_g}{\|\mathbf{v}_{gi}\|\|\mathbf{v}_g\|} = \frac{\sum_{k=1}^{K_g} t_{gik}t_{gk}}{\sqrt{\sum_{k=1}^{K_g} t_{gik}^2}\sqrt{\sum_{k=1}^{K_g} t_{gk}^2}}$ is the cosine similarity of the sentence vector $\mathbf{v}_{gi}$ with its good mean $\mathbf{v}_g$,[80] K_g is the number of tokens used in descriptions of good g, and $\bar{s}_g = \sum_{i=1}^{|\mathbf{t}_g|} s\left(\mathbf{v}_{gi}, \mathbf{v}_g\right)$ is the mean of good g's sentence cosine similarities. We deemed sentences to be correctly classified if their predicted value $\hat{g}_C\left(\mathbf{x}_i\right)$ was above the median and their normalized cosine similarity θ_{gi} was above the median. Figure D.1 panels A and B and Table E.4 show the robustness of our results to using the 45th or 55th percentile as thresholds.

A3 Clustering

The third step of our procedure takes the misclassified sentences from the classification step and groups them into clusters of similar sentences. We will then use these clusters as our good classification for this group of purchases. To per-

[80]Note that the cosine similarity ranges from 0 to 1, with 0 being orthogonal vectors and 1 indicating vectors pointing in the same direction.

form this clustering we use the popular K-means method. This method groups the tokenized sentences into k clusters by finding a centroid c_k for each cluster to minimize the sum of squared distances between the sentences and their group's centroid. That is, it solves

$$\min_{\mathbf{c}} \sum_{i=1}^{N} \left\| f\left(\mathbf{c}, \mathbf{t}_i\right) - \mathbf{t}_i \right\|^2 \tag{A.5}$$

where $f(\mathbf{c}, \mathbf{t}_i)$ returns the closest centroid to $\mathbf{t}_i$. To speed up the clustering on our large dataset we implemented the algorithm by mini-batch k-means. Mini-batch k means iterates over random subsamples (in our case of size 500) to minimize computation time. In each iteration, each sentence is assigned to it's closest centroid, and then the centroids are updated by taking a convex combination of the sentence and its centroid, with a weight on the sentence that converges to zero as the algorithm progresses (see Sculley (2010) for details).

The key parameter choice for the clustering exercise is k, the number of clusters to group the sentences into. As is common in the literature, we make this choice using the silhouette coefficient. For each sentence, its silhouette coefficient is given by

$$\eta(i) = \frac{b(i) - a(i)}{\max\{b(i), a(i)\}} \tag{A.6}$$

where $a(i)$ is the average distance between sentence i and the other sentences in the same cluster, and $b(i)$ is the average distance between sentence i and the sentences in the nearest cluster to sentence i's cluster. A high value of the silhouette coefficient indicates that the sentence is well clustered: it is close to the sentences in its cluster and far from the sentences in the nearest cluster. We start by using a k of 300 for each 2-digit product categories. For 2-digit product categories with an average silhouette coefficient larger than the overall average silhouette coefficient, we tried $k \in \{250, 200, 150, 100, 50, 25, 10, 7\}$ while for product categories with a lower than average silhouette coefficient we tried $k \in \{350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000\}$ until the average silhouette score was equalized across 2-digit product codes.

Proofs of Propositions

B1 Proof of Proposition 1

PROOF:

The suppliers choose their entry and bidding strategies to maximize expected profits. 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; she receives the contract at the other bidder's fulfillment cost. The expected profits from an auction in which firm i bids b_i are then $\mathbb{E}\left[\pi_i|b_i\right] = \mathbb{E}_{b_j}\left[b_j - b_i|b_j > b_i\right]\mathbb{P}\left(b_j > b_i\right)$ making the expected profits from the auction to bidder i, $\mathbb{E}\left[\pi_i\right] = \mathbb{E}_{b_i}\left[\mathbb{E}\left[\pi_i|b_i\right]\right]$.

Working back to the entry decisions, the two firms enter with probabilities q_F and q_L. If firm i pays the participation cost c_i and enters, with probability q_j firm j also enters and the auction takes place, yielding firm i expected profits of $\mathbb{E}\left[\pi_i\right]$, while with probability $1-q_j$, i is the only entrant and receives the contract at price $\bar{\theta}$ yielding expected profits of $\bar{\theta} - \mathbb{E}\left[\bar{\theta}/\theta_i\right]$. If instead firm i chooses not to enter, her profits are zero but she does not have to pay the participation cost. The nature of the equilibrium depends on the size of the participation costs c_i. When participation costs are sufficiently small, both firms enter with certainty and the auction always takes place. For larger participation costs the equilibrium involves mixed strategies. In a mixed strategy equilibrium, the firms are indifferent between entering and not entering, pinning down the entry probabilities

$$q_j\mathbb{E}\left[\pi_i\right] + (1-q_j)\left(\bar{\theta} - \mathbb{E}\left[\bar{\theta}/\theta_i\right]\right) = c_i \iff q_j = \frac{\bar{\theta} - \mathbb{E}\left[\bar{\theta}/\theta_i\right] - c_i}{\bar{\theta} - \mathbb{E}\left[\bar{\theta}/\theta_i\right] - \mathbb{E}\left[\pi_i\right]}, \tag{B.1}$$

where $i, j \in \{F, L\}$, $i \neq j$.

For the firms to be indifferent between entering and not entering, equation (B.1) must hold. Solving the equation requires us to derive expressions for $\mathbb{E}\left[b_i\right]$ and $\mathbb{E}\left[\pi_i\right]$. The distribution of the bids is given by the bidding functions $b_i = \bar{\theta}/\theta_i$ and the Pareto distributions of the productivities θ_i: $G_i\left(\theta_i\right) = 1 - \theta_i^{-\delta_i}$.

$$H_i\left(b\right) \equiv \mathbb{P}\left(b_i \leq b\right) = \mathbb{P}\left(\theta_i \geq \frac{\bar{\theta}}{b}\right) = \left(\frac{b}{\bar{\theta}}\right)^{\delta_i} \tag{B.2}$$

The expected bids are then simply $\mathbb{E}\left[b_i\right] = \int_0^{\bar{\theta}} b dH_i\left(b\right) = \frac{\delta_i}{1+\delta_i}\bar{\theta}$.

To derive expected profits from the auction $\mathbb{E}\left[\pi_i\right]$ we begin by considering expected profits conditional on a bidders fulfillment cost. Since the optimal bidding strategies are to bid the firm's true valuation, expected profits for a firm with valuation b_i are

$$\mathbb{E}\left[\pi_i|b_i\right] = \mathbb{E}_{b_j}\left[b_j - b_i|b_j > b_i\right]\mathbb{P}\left(b_j > b_i\right) = \int_{b_i}^{\bar{\theta}}\left(b_j - b_i\right)dH_j\left(b_j\right)$$

$$= \frac{\delta_j}{1+\delta_j}\bar{\theta} - b_i + b_i\left(\frac{b_i}{\bar{\theta}}\right)^{\delta_j}\frac{1}{1+\delta_j}, \tag{B.3}$$

where the final equality follows by inserting (B.2) and integrating. Now we can derive unconditional expected profits by the law of iterated expectations:

$$\mathbb{E}\left[\pi_i\right] = \mathbb{E}_{b_i}\left[\mathbb{E}\left[\pi_i|b_i\right]\right] = \int_0^{\bar{\theta}}\mathbb{E}\left[\pi_i|b_i\right]dH_i\left(b_i\right) = \left(\frac{1}{1+\delta_i} - \frac{1}{1+\delta_F+\delta_L}\right)\bar{\theta}. \tag{B.4}$$

Inserting these and the definition of the entry costs c_i into (B.1) and rearranging yields the statement in the proposition

$$q_i = \sqrt{\kappa\left(1-\alpha_c-\psi_c\right)}, \tag{B.5}$$

where $\kappa = \min\left\{\left[\left(1+\delta_F+\delta_L\right)/\left(1+\delta_L\right)\right]^2, 1/\left(1-\alpha_c-\psi_c\right)\right\}$.

Turning to the expected prices, whenever neither or only one firm enters, the price is $\bar{\theta}$. When both enter, the price is the higher of the two bids.

$$\mathbb{P}\left(p \leq x\right) = \mathbb{P}\left(\max\{b_F, b_L\} \leq x\right) = H_F\left(x\right)H_L\left(x\right) = \left(\frac{x}{\bar{\theta}}\right)^{\delta_F+\delta_L} \tag{B.6}$$

As a result, the distribution and expectation of the log price when both firms enter is

$$\mathbb{P}\left(\log\left(p\right) \leq x\right) = \mathbb{P}\left(p \leq \mathrm{e}^x\right) = \left(\frac{\mathrm{e}^x}{\bar{\theta}}\right)^{\delta_F+\delta_L}$$

(B.7)

$$\mathbb{E}\left[\log\left(p\right) \mid \text{both enter}\right] = \int_{-\infty}^{\log(\bar{\theta})} x\frac{\delta_F+\delta_L}{M^{\delta_F+\delta_L}}e^{(\delta_F+\delta_L)x}dx = \log\left(\bar{\theta}\right) - \frac{1}{\delta_F+\delta_L}$$

The expected log price is then simply $\mathbb{E}\left[\log\left(p\right)\right] = q_F q_L \mathbb{E}\left[\log\left(p\right) \mid \text{both enter}\right] + \left(1 - q_F q_L\right)\log\left(\bar{\theta}\right)$. Inserting (B.7) and the entry probabilities q_F and q_L yields expression (1) in the proposition.

The comparative statics on prices follow straightforwardly from equation (1). The comparative static on the number of bidders follows straightforwardly from noting that the expected number of entrants is $q_F + q_L$.

B2 Proof of Proposition 2

PROOF:

In this setting it is optimal for bidder F to shade so that her bid net of the bid penalty is equal to her true 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. This means that for any given bid, the preference regime lowers expected profits for foreign bidders and increases them for local bidders, as the policy intends. To see this, note that the expected profits of bids b_F and b_L are now

(B.8)

$$\mathbb{E}\left[\pi_F|b_F,\gamma\right] = \mathbb{E}\left[\gamma\left(b_L - b_F\right)|b_L > b_F\right]\mathbb{P}\left(b_L > b_F\right)$$
$$\mathbb{E}\left[\pi_L|b_L,\gamma\right] = \mathbb{E}\left[b_F - b_L|\bar{\theta} \geq b_F > b_L\right]\mathbb{P}\left(\bar{\theta} \geq b_F > b_L\right) + \mathbb{P}\left(\theta_F < 1/\gamma\right)\left(\bar{\theta} - b_L\right).$$

For any particular bid, the profits to bidder F are shrunk by the penalty γ, forcing bidder F to bid more aggressively and lowering expected profits. For bidder L the probability of winning with any bid increases, and the bid penalty creates a discrete probability that bidder F drops out, both of which increase L's expected profits.

Consider the three cases in proposition 2 in turn.

Buyers with $\alpha_c + \psi_c \leq \underline{c}$.

In this case, both bidders enter the auction with certainty. Entering the auction is a best response to the other bidder entering whenever $\mathbb{E}\left[\pi_i|\gamma\right] - c_i > 0$. Expected profits are lower for bidder F and participation costs c_F are higher, so bidder F is the pivotal bidder for this case. Integrating bidder F's expected profits conditional on her bid (B.8) over all bids,

(B.9)

$$\mathbb{E}\left[\pi_F|\gamma < 1\right] = \int_0^M \mathbb{E}\left[\pi_F|b_F,\gamma < 1\right]dH_F\left(b_F|\gamma < 1\right) = \gamma^{1+\delta_F}M\left(\frac{1}{1+\delta_F} - \frac{1}{1+\delta_F+\delta_L}\right)$$

Setting (B.9) equal to c_F and rearranging yields the definition of $\underline{c}$ in the proposition. Since $\underline{c} < 1 - \left(\frac{1+\delta_L}{1+\delta_F+\delta_L}\right)^2$, both bidders enter the auction with or without the preferences and so participation is unchanged.

Since bidding behavior has changed, the expected price in the auction has changed. There are three possibilities:

$$p = \begin{cases} b_F & \text{if } b_L < b_F < \bar{\theta}, \\ \bar{\theta} & \text{if } b_L < M \leq b_F, \\ \gamma b_L & \text{if } b_F \leq b_L. \end{cases}$$

Combining these the distribution of prices is given by

$$\mathbb{P}\left(p \leq x\right) = \begin{cases} H_F\left(x\right) H_L\left(x/\gamma\right) + \int_x^{x/\gamma} \int_{b_F}^{x/\gamma} h_L\left(b_L\right) db_L h_F\left(b_F\right) db_F & \text{if } 0 \leq x \leq \gamma\bar{\theta}, \\ H_F\left(x\right) + \int_x^{\bar{\theta}} \int_{b_F}^{\bar{\theta}} h_L\left(b_L\right) db_L h_F\left(b_F\right) db_F & \text{if } \gamma\bar{\theta} < x < \bar{\theta}, \\ 1 & \text{if } x = \bar{\theta} \end{cases}$$

$$= \begin{cases} \left(\frac{\delta_L}{\delta_F+\delta_L}\gamma^{-\delta_F-\delta_L} + \frac{\delta_F}{\delta_F+\delta_L}\right) H_F\left(x\right) H_L\left(x\right) & \text{if } 0 \leq x \leq \gamma\bar{\theta}, \\ \frac{\delta_L}{\delta_F+\delta_L}\gamma^{\delta_F} + \frac{\delta_F}{\delta_F+\delta_L} H_F\left(x\right) H_L\left(x\right) & \text{if } \gamma\bar{\theta} < x < \bar{\theta}, \\ 1 & \text{if } x = \bar{\theta} \end{cases}$$

In turn, the distribution of log prices is given by

$$\mathbb{P}\left(\log\left(p\right) \leq x\right) = \mathbb{P}\left(p \leq \mathrm{e}^x\right) = \begin{cases} \left(\frac{\delta_L}{\delta_F+\delta_L}\gamma^{-\delta_L} + \frac{\delta_F}{\delta_F+\delta_L}\gamma^{\delta_F}\right)\left(\frac{\mathrm{e}^x}{\bar{\theta}}\right)^{\delta_F+\delta_L} & \text{if } -\infty < x \leq \log\left(\gamma\bar{\theta}\right), \\ \frac{\delta_L}{\delta_F+\delta_L}\gamma^{\delta_F} + \frac{\delta_F}{\delta_F+\delta_L}\gamma^{\delta_F}\left(\frac{\mathrm{e}^x}{\bar{\theta}}\right)^{\delta_F+\delta_L} & \text{if } \log\left(\gamma\bar{\theta}\right) < x < \log\left(\bar{\theta}\right), \\ 1 & \text{if } x = \log\left(\bar{\theta}\right) \end{cases}$$

making the expected log price in the auction

$$\begin{aligned} \mathbb{E}\left[\log\left(p\right)|\text{both enter}\right] &= \int_{-\infty}^{\log\left(\gamma\bar{\theta}\right)} \frac{\delta_L\gamma^{-\delta_L}+\delta_F\gamma^{\delta_F}}{\bar{\theta}^{\delta_F+\delta_L}} x\mathrm{e}^{(\delta_F+\delta_L)x}\mathrm{d}x + \int_{\log\left(\gamma\bar{\theta}\right)}^{\log\left(\bar{\theta}\right)} \frac{\delta_F\gamma^{\delta_F}}{\bar{\theta}^{\delta_F+\delta_L}} x\mathrm{e}^{(\delta_F+\delta_L)x}\mathrm{d}x \\ &\quad + \left[1 - H_F\left(\bar{\theta}\right)\right]\log\left(\bar{\theta}\right) \\ &= \log\left(\bar{\theta}\right) - \frac{\gamma^{\delta_F}\left(1-\log\left(\gamma^{\delta_L}\right)\right)}{\delta_F+\delta_L}. \end{aligned} \tag{B.10}$$

Comparing (B.10) to the expected price without preferences (B.7), prices rise as long as $\gamma^{\delta_F}\left[1-\log\left(\gamma^{\delta_L}\right)\right] < 1$.

Finally, the probability that the local bidder wins the auction when there are no preferences is

$$\mathbb{P}\left(L \text{ wins}\right) = \mathbb{P}\left(b_L < b_F\right) = \int_0^{\bar{\theta}} H_L\left(b_F|\gamma=1\right) dH_F\left(b_F|\gamma=1\right) = 1 - \frac{\delta_L}{\delta_F+\delta_L}, \tag{B.11}$$

while when there are preferences this increases to

$$\mathbb{P}\left(L \text{ wins}\right) = \mathbb{P}\left(b_L < b_F|\gamma<1\right) = \int_0^{\bar{\theta}} H_L\left(b_F|\gamma<1\right) dH_F\left(b_F|\gamma<1\right) = 1-\gamma^{\delta_F}\frac{\delta_L}{\delta_F+\delta_L}. \tag{B.12}$$

BUYERS WITH $\underline{c} < \alpha_c + \psi_c \leq \bar{c}$.

This case occurs when bidder L finds it worthwhile to enter the auction with certainty and bidder F's best response is to remain out of the auction with certainty. That is, when $\mathbb{E}[\pi_F|\underline{\gamma}] - c_F < 0$ and $\mathbb{E}[\pi_L|\gamma] - c_L > 0$. In this case, since only L enters, the price is $\bar{\theta}$ with certainty, which is higher than in the absence of preferences since in the absence of preferences the auction always takes place with positive probability. Participation is therefore also lower, and since bidder L now wins with certainty, the probability that bidder L wins has increased.

The threshold $\underline{c}$ is defined in the previous case as the solution to $\mathbb{E}[\pi_L|\gamma] - c_L = 0$. To find the upper threshold $\bar{c}$, we require an expression for $\mathbb{E}[\pi_L|\gamma]$:

$$\mathbb{E}[\pi_L|\gamma < 1] = \int_0^{\bar{\theta}} \mathbb{E}[\pi_L|b_L, \gamma < 1]\, dH_L(b_L|\gamma < 1) = \bar{\theta}\left(\frac{1}{1+\delta_L} - \frac{\gamma^{\delta_F}}{1+\delta_F+\delta_L}\right). \tag{B.13}$$

Setting (B.13) equal to c_L and rearranging yields the definition of $\underline{c}$ in the proposition.

Buyers with $\bar{c} < \alpha_c + \psi_c$.

This case occurs when neither bidder finds it optimal to enter with certainty: $\mathbb{E}[\pi_i|\gamma] - c_i < 0\ \forall i$ and so the equilibrium is in mixed strategies. As in proposition 1, the entry probabilities are given by

$$q_i = \frac{\bar{\theta} - \mathbb{E}\left[\bar{\theta}/\theta_j\right] - c_j}{\bar{\theta} - \mathbb{E}[b_j] - \mathbb{E}[\pi_j|\gamma < 1]}.$$

In this case the expected price is given by

$$\mathbb{E}[\log(p)] = \log\left(\bar{\theta}\right) - q_F q_L \left(\log\left(\bar{\theta}\right) - \mathbb{E}[\log(p)\,|\text{both enter}]\right)$$

Inserting the entry probabilities and the price equation (B.10) and rearranging, the expected price when there are preferences is lower whenever

$$\begin{aligned} &q_F(\gamma < 1)\, q_L(\gamma < 1)\left(\log\left(\bar{\theta}\right) - \mathbb{E}[\log(p)\,|\text{both enter}, \gamma < 1]\right) \\ &-q_F(\gamma = 1) q_L(\gamma = 1)\left(\log\left(\bar{\theta}\right) - \mathbb{E}[\log(p)\,|\text{both enter}, \gamma = 1]\right) \geq 0 \\ &\iff -\log\left(\gamma^{\delta_L}\right) - \frac{\delta_L}{1+\delta_F}\left(1 - \gamma^{1+\delta_F}\right) \geq 0 \end{aligned} \tag{B.14}$$

Noting that (B.14) holds with equality when $\gamma = 1$ and that the left hand side of (B.14) has slope $-\delta_L\left(\gamma^{-1} - \gamma^{\delta_F}\right) < 0\ \forall \gamma < 1$ shows that (B.14) holds for all $\gamma < 1$. Participation in the auction is $\mathbb{E}[N] = q_F + q_L$. When there are no preferences

$$\mathbb{E}[N|\gamma = 1] = q_F(\gamma = 1) + q_L(\gamma = 1) = 2\frac{1+\delta_F+\delta_L}{1+\delta_L}\sqrt{1-\alpha_c-\psi_c}, \tag{B.15}$$

while with preferences participation is

$$\mathbb{E}\left[N|\gamma<1\right] = q_F\left(\gamma<1\right)+q_L\left(\gamma<1\right)$$
$$= \left(\frac{1}{\gamma^{\delta_F}}+\frac{1}{\gamma^{1+\delta_F}+\left(1-\gamma^{1+\delta_F}\right)\frac{1+\delta_F+\delta_L}{1+\delta_F}}\right)\frac{1+\delta_F+\delta_L}{1+\delta_L}\sqrt{1-\alpha_c-\psi_c}. \tag{B.16}$$

Comparing (B.15) to (B.16) shows that participation increases whenever

$$\frac{1}{\gamma^{\delta_F}}+\frac{1}{1+\frac{\delta_L}{\delta_F+\delta_L}\left(1-\gamma^{1+\delta_F}\right)}>2 \tag{B.17}$$

Equation (B.17) is implied by our assumption that we are in the case where $\gamma^{\delta_F}\left[1+\frac{\delta_L}{\delta_F+\delta_L}\left(1-\gamma^{1+\delta_F}\right)\right]<1$

Finally, to see that the probability that bidder L wins the contract at auction increases by more than in case 1 note that the probability that bidder L wins the contract is given by $q_F q_L \mathbb{P}\left(b_F<b_L\right)$. The probability that bidder L wins will increase by more if $q_F\left(\gamma=1\right)q_L\left(\gamma=1\right)<q_F\left(\gamma<1\right)q_L\left(\gamma<1\right)$. Computing the components of this

$$\frac{q_F\left(\gamma=1\right)}{q_F\left(\gamma<1\right)}=\frac{\bar{\theta}-\mathbb{E}\left[\bar{\theta}/\theta_L\right]-\mathbb{E}\left[\pi_L|\gamma<1\right]}{\bar{\theta}-\mathbb{E}\left[\bar{\theta}/\theta_L\right]-\mathbb{E}\left[\pi_L|\gamma=1\right]}=\gamma^{\delta_F}$$
$$\frac{q_L\left(\gamma=1\right)}{q_L\left(\gamma<1\right)}=\frac{\bar{\theta}-\mathbb{E}\left[\bar{\theta}/\theta_F\right]-\mathbb{E}\left[\pi_F|\gamma<1\right]}{\bar{\theta}-\mathbb{E}\left[\bar{\theta}/\theta_F\right]-\mathbb{E}\left[\pi_F|\gamma=1\right]}=1+\frac{\delta_L}{1+\delta_F}\left(1-\gamma^{1+\delta_F}\right)$$

Combining these two components shows that the statement is correct as long as

$$\gamma^{-\delta_F}>\left[1+\frac{\delta_L}{\delta_F+\delta_L}\left(1-\gamma^{1+\delta_F}\right)\right] \tag{B.18}$$

Condition is implied by the condition stated at the top of the proposition that $\gamma^{-\delta_F}>1-\log\left(\gamma^{\delta_L}\right)$. To see this, note that both conditions are decreasing in γ and that their limits are the same as γ approaches 1 from below. Then, note that the slope of the right-hand-side of the condition in the proposition is steeper than the slope of condition (B.18): The slope of the condition in the proposition is $-\delta_L\gamma^{-1}$ while the slope of condition (B.18) is $-\frac{\delta_F}{\delta_L+\delta_F}\left(1+\delta_F\right)\gamma^{\delta_F}$ which is flatter since rearranging

$$-\delta_L\gamma^{-1}<-\frac{\delta_F}{\delta_L+\delta_F}\left(1+\delta_F\right)\gamma^{\delta_F}\iff\frac{\delta_L}{\delta_F}\frac{\delta_L+\delta_F}{1+\delta_F}>\gamma^{1+\delta_F} \tag{B.19}$$

and both terms on the left are larger than 1 while the term on the right is smaller than 1. Hence, the condition in the proposition implies condition (B.18).

C: Identification of Bureaucrat and Organization Effects with Multiple Connected Sets

As shown in Abowd, Creecy and Kramarz (2002), it isn't possible to identify all the bureaucrat and organization effects. In particular, they show that (a) the effects are identified only within connected sets of bureaucrats and organizations; and (b) within each connected set s containing $N_{b,s}$ bureaucrats and $N_{o,s}$ organizations, only the group mean of the lhs variable, and $N_{b,s} - 1 + N_{o,s} - 1$ of the bureaucrat and organization effects are identified. More generally, within each connected set, we can identify $N_{b,s} + N_{o,s} - 1$ linear combinations of the bureaucrat and organization effects.

To see this explicitly, write the model as

$$\mathbf{p} = \mathbf{X}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\alpha} + \mathbf{F}\boldsymbol{\psi} \tag{C.1}$$

where $\mathbf{p}$ is the $N \times 1$ vector of item prices; $\mathbf{X}$ is an $N \times k$ matrix of control variables, $\mathbf{B}$ is the $N \times N_b$ design matrix indicating the bureaucrat responsible for each purchase; $\boldsymbol{\alpha}$ is the $N_b \times 1$ vector of bureaucrat effects; $\mathbf{F}$ is the $N \times N_o$ design matrix indicating the organization responsible for each purchase; and $\boldsymbol{\psi}$ is the $N_o \times 1$ vector of organization effects.

Suppressing $\mathbf{X}\boldsymbol{\beta}$ for simplicity, the OLS normal equations for this model are

$$\begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\alpha}}_{OLS} \\ \hat{\boldsymbol{\psi}}_{OLS} \end{bmatrix} = \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \end{bmatrix} \mathbf{p} \tag{C.2}$$

As Abowd, Creecy and Kramarz (2002) show, these equations do not have a unique solution because $[\mathbf{B}\,\mathbf{F}]'[\mathbf{B}\,\mathbf{F}]$ only has rank $N_b + N_o - N_s$, where N_s is the number of connected sets. As a result, to identify a particular solution to the normal equations, we need N_s additional restrictions on the αs and ψs.

Abowd, Creecy and Kramarz (2002) add N_s restrictions setting the mean of the person effects to 0 in each connected set. They also set the grand mean of the firm effects to 0. However, this makes it difficult to compare across connected sets since all the firm effects are interpreted as deviations from the grand mean, which is a mean across connected sets. Instead, we will add $2N_s$ restrictions setting the mean of the bureaucrat and organization effects to 0 within each connected set. These N_s additional constraints also allow us to identify S connected set means $\gamma_s = \bar{\alpha}_s + \bar{\psi}_s$ which facilitate comparison across connected sets and allow us to interpret the variances of the estimated bureaucrat and organization effects as lower bounds on the true variances of the bureaucrat and organization effects.

Specifically, we augment the model to be

$$\mathbf{p} = \mathbf{B}\tilde{\boldsymbol{\alpha}} + \mathbf{F}\tilde{\boldsymbol{\psi}} + \mathbf{S}\boldsymbol{\gamma} \tag{C.3}$$

where $\mathbf{S}$ is the $N \times N_s$ design matrix indicating which connected set each item belongs to; $\boldsymbol{\gamma}$ is the $N_s \times 1$ vector of connected set effects; and we add the restriction that $\tilde{\alpha}$ and $\tilde{\psi}$ have mean zero in each connected set. Our fixed effects estimates thus solve the normal equations of this augmented model, plus $2N_s$ zero-mean restrictions:

$$\text{(C.4)} \qquad \begin{bmatrix} \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \\ \mathbf{S}' \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{F} & \mathbf{S} \end{bmatrix} \\ \begin{bmatrix} \mathbf{S_b} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S_o} & \mathbf{0} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\psi}} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \\ \mathbf{S}' \end{bmatrix} \mathbf{p} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

where $\mathbf{S}_b$ is the $N_s \times N_b$ design matrix indicating which connected set each bureaucrat belongs to, and $\mathbf{S}_o$ is the $N_s \times N_o$ design matrix indicating which connected set each organization belongs to.

The following proposition describes the relationship between these estimators and the bureaucrat and organization effects.

PROPOSITION 3 (**Identification**): *If the true model is given by* (C.1)*, then* $\hat{\boldsymbol{\alpha}}$*,* $\hat{\boldsymbol{\psi}}$*, and* $\hat{\boldsymbol{\gamma}}$*, the estimators of* $\tilde{\boldsymbol{\alpha}}$*,* $\tilde{\boldsymbol{\psi}}$ *and* $\boldsymbol{\gamma}$ *in the augmented model* (C.3) *that solve the augmented normal equations* (C.4) *(i) are uniquely identified, and (ii) are related to the true bureaucrat and organization effects* $\boldsymbol{\alpha}$ *and* $\boldsymbol{\psi}$ *by*

$$\text{(C.5)} \qquad \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\psi}} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha} - \mathbf{S_b}'\overline{\boldsymbol{\alpha}} \\ \boldsymbol{\psi} - \mathbf{S_o}'\overline{\boldsymbol{\psi}} \\ \overline{\boldsymbol{\alpha}} + \overline{\boldsymbol{\psi}} \end{bmatrix}$$

where $\overline{\boldsymbol{\alpha}}$ *is the* $N_s \times 1$ *vector of connected-set bureaucrat effect means, and* $\overline{\boldsymbol{\psi}}$ *is the* $N_s \times 1$ *vector of connected-set organization effect means.*

PROOF:

We will prove each part of the result separately. To see uniqueness, first note that the standard normal equations for (C.3) only has rank $N_b + N_o - N_s$. To see this, we note that $\mathbf{B}\mathbf{S_b}' = \mathbf{F}\mathbf{S_o}' = \mathbf{S}$ and so $2N_s$ columns of the $N \times (N_b + N_o + N_s)$ matrix $[\mathbf{B}\ \mathbf{F}\ \mathbf{S}]$ are collinear. However, the $2N_s$ restrictions $\mathbf{S_b}\hat{\boldsymbol{\alpha}} = 0$ and $\mathbf{S_o}\hat{\boldsymbol{\psi}} = 0$ are independent of the standard normal equations, so the first matrix in (C.4) has rank $N_b + N_o + N_s$ and hence the solution to (C.4) is unique.

To see the second part, it suffices to show that (C.5) solves (C.4). First, substitute the estimators out of (C.4) using (C.5) and substitute in the true model using (C.1) to rewrite (C.4) as

$$\begin{bmatrix} \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \\ \mathbf{S}' \end{bmatrix} \left[\mathbf{B}\left(\boldsymbol{\alpha} - \mathbf{S_b}'\overline{\boldsymbol{\alpha}}\right) + \mathbf{F}\left(\boldsymbol{\psi} - \mathbf{S_o}'\overline{\boldsymbol{\psi}}\right) + \mathbf{S}\left(\overline{\boldsymbol{\alpha}} + \overline{\boldsymbol{\psi}}\right)\right] \\ \mathbf{S_b}\left(\boldsymbol{\alpha} - \mathbf{S_b}'\overline{\boldsymbol{\alpha}}\right) \\ \mathbf{S_o}\left(\boldsymbol{\psi} - \mathbf{S_o}'\overline{\boldsymbol{\psi}}\right) \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \\ \mathbf{S}' \end{bmatrix} [\mathbf{B}\boldsymbol{\alpha} + \mathbf{F}\boldsymbol{\psi}] \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

From here, noting again that $\mathbf{B}\mathbf{S_b}' = \mathbf{F}\mathbf{S_o}' = \mathbf{S}$; that $\mathbf{S_b}\boldsymbol{\alpha}$ is an $N_s \times 1$ vector in which each entry is the sum of the bureaucrat effects; and that $\mathbf{S_o}\boldsymbol{\psi}$ is an $N_s \times 1$ vector in which each entry is the sum of the organization effects, shows that the two sides are equal, yielding the result.

The above analysis focuses on the simple case in which there are no other covariates in the model. In the more general model with covariates it is not always possible to separately identify the connected set intercepts $\boldsymbol{\gamma}$, particularly when the covariates $\mathbf{X}$ include categorical variables. Nevertheless, the identification of the bureaucrat effects $\tilde{\boldsymbol{\alpha}}$ and organization effects $\tilde{\boldsymbol{\psi}}$ remains as above. In our empirical application we have categorical covariates and so we focus on the bureaucrat- and organization- effects and do not results on the connected set intercepts $\boldsymbol{\gamma}$.

D: Additional Results on Event Studies to Identify the Effectiveness of Individuals and Organizations

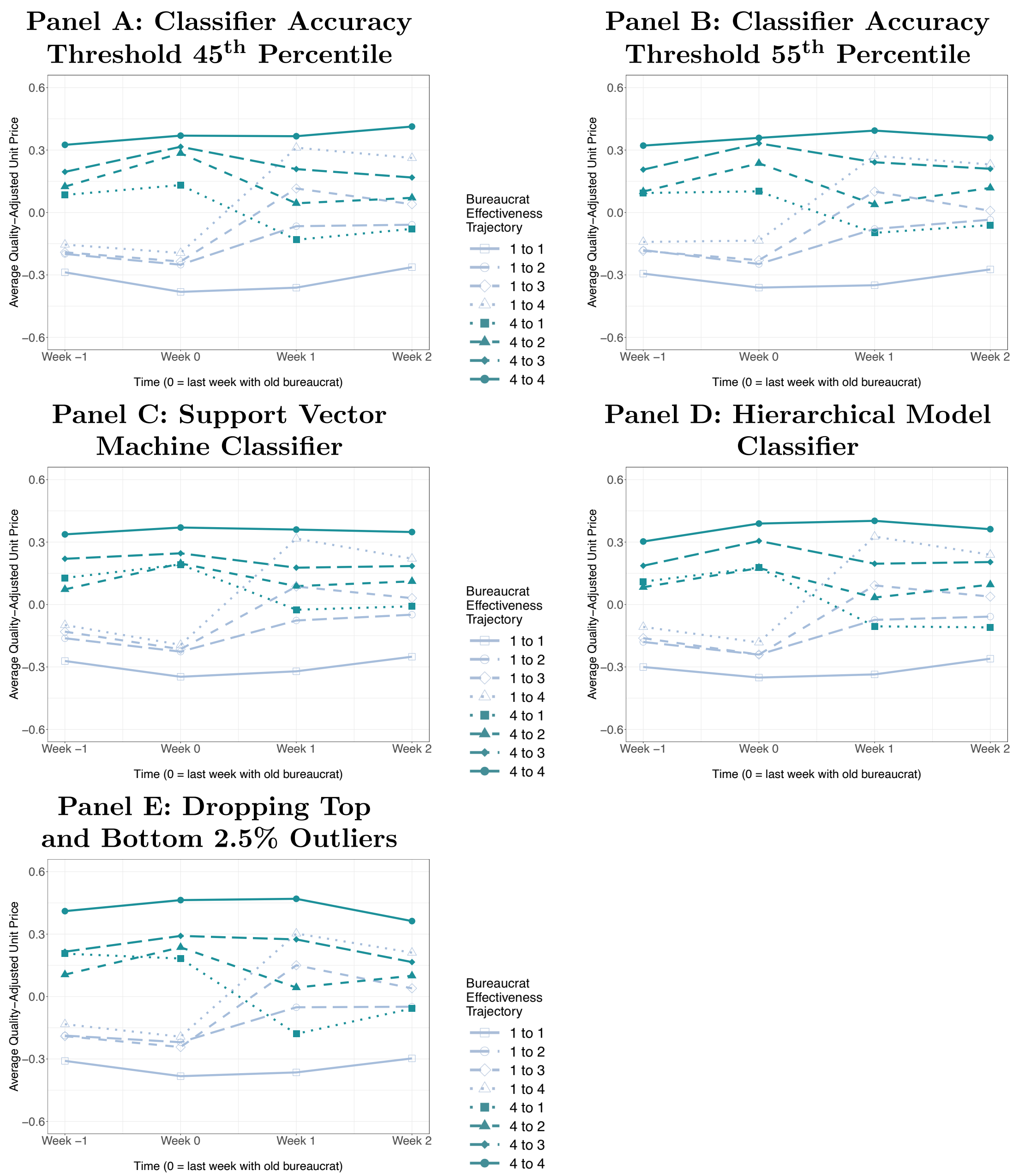
In Sub-section IV.A we argue that using event studies around the time that organizations change the bureaucrat they work with can identify their effectiveness. In this appendix, we show that this argument is robust to changing a series of choices made in constructing the event studies.

In figure D.1 we consider the choices made in how the sample was built for the analysis. As described in Appendix A, we deemed contract descriptions to be correctly classified whenever their predicted value and their normalized cosine similarity with their labeled good's mean vector were both above the median. In Panel A we instead classify them as correct whenever they are above the 45th percentile, and in panel B we use the 55th percentile as our threshold. The results are essentially unchanged. Our baseline classifier uses the logistic function as its objective function, which performs very well. Nevertheless, in Panel C, we instead use a support vector machine (SVM) objective function. And in Panel D we train a sequence of hierarchical classifiers exploiting the hierarchical structure of the HS product codes (details are in Appendix A). In both cases, the results are unchanged. Finally, in panel E, we trim the top and bottom 2.5% of each product rather than the 5% we use in our baseline data. Again, the results are unaffected.

In Figures D.2 and D.3 we change a series of the choices made in constructing the event studies. In figure D.2 we vary the units of time we use to define the spells that we combine to create events. In Figure 1 we define a spell as a sequence of two *weeks*, separated by fewer than 400 days. In Panel A, rather than weeks, we use days. In Panel B we use fotnights. In panel C we use months. And in Panel D we define a spell as a sequence of *three* weeks instead of two. The results are very similar in all cases. In figure D.3 we consider four more design choices. In Panel A we use a coarser categorization of the effectiveness of the bureaucrats in each event: We use terciles instead of quartiles. In Panel B we use a global ranking of bureaucrats (instead of a separate ranking for each semester as in Figure 1). Finally, we consider spells in which the weeks are separated by up to 350 days (Panel C) or 450 days (Panel D) rather than 400 days. In all cases, the results are very similar.

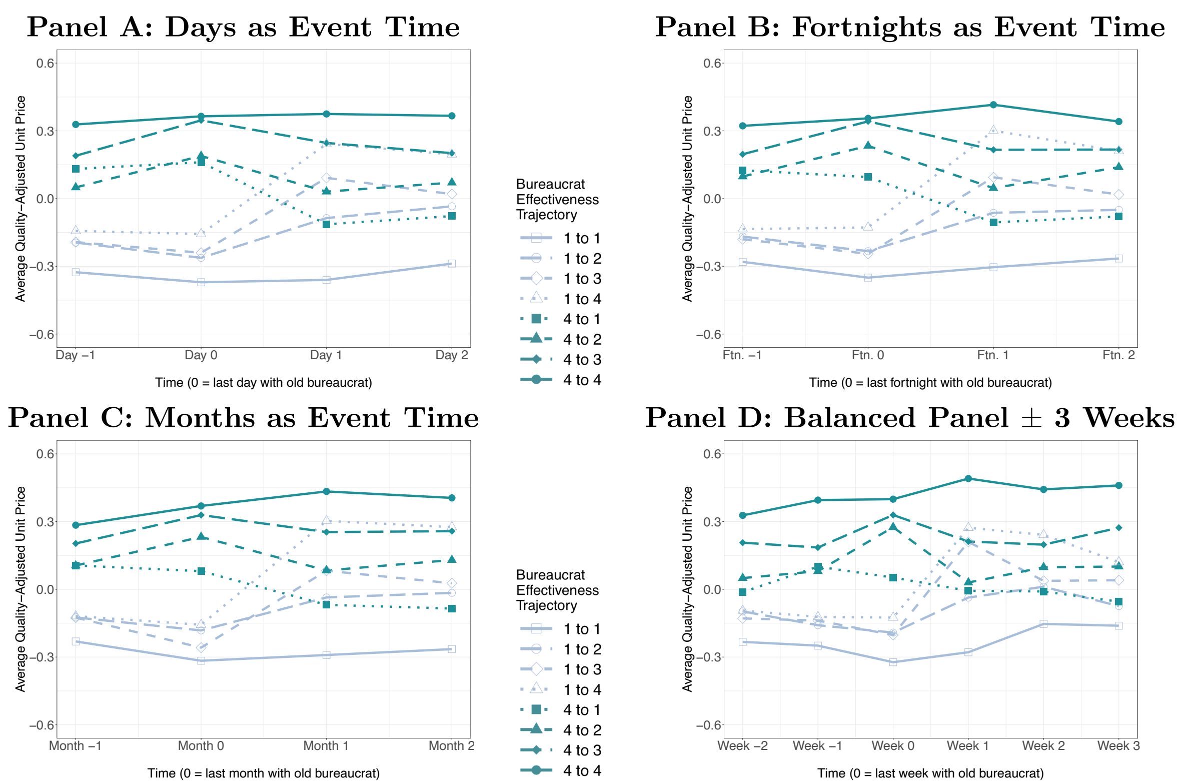
Our main event study studies the prices paid by organizations around the time they switch the bureaucrat they work with. Figure D.4 considers two other such changes: Bureaucrats switching which good they buy (Panel A) and organizations switching which good they buy (Panel B). Again, the results strongly support the use of switches to identify the effectiveness of bureaucrats and organizations. Table D.1 displays the data underlying our main event study in Figure 1 along with some additional summary statistics on the event study (the sample sizes in columns (1) and (2) and the time gaps between event time periods in columns (7)–(9)). Table D.2 compares the sample used in the event study to the analysis sample, showing that the two samples are comparable.

Figure D.1. Robustness of Event Studies to Alternative Text Classifiers, Classification Accuracy Thresholds, and Outlier Trimming



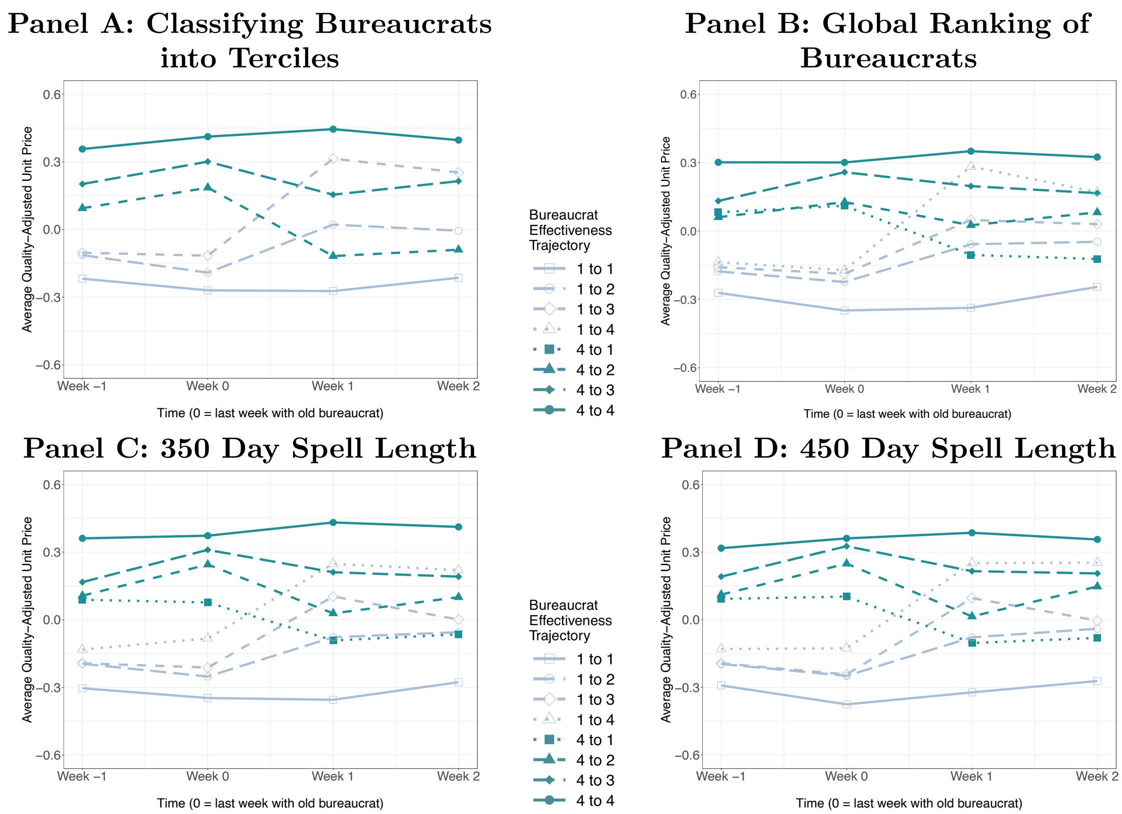
Note: Each panel in the figure is analogous to Figure 1 (see notes to that figure for details of construction), with the following changes. Rather than requiring our classifier's predicted value and normalized cosine similarity to be above the median, we require them to be above the 45th percentile (Panel A) or the 55th percentile (Panel B). The classifier in Panel C uses a support vector machine objective function rather than a logistic function. The classifier in Panel D is a hierarchical series of classifiers exploiting the hierarchical structure of HS codes. See appendix A for details. Finally, in Panel E, we trim the top and bottom 2.5% of each product rather than the 5% we use in our baseline data.

FIGURE D.2. ROBUSTNESS OF EVENT STUDIES TO DESIGN CHOICES (1)



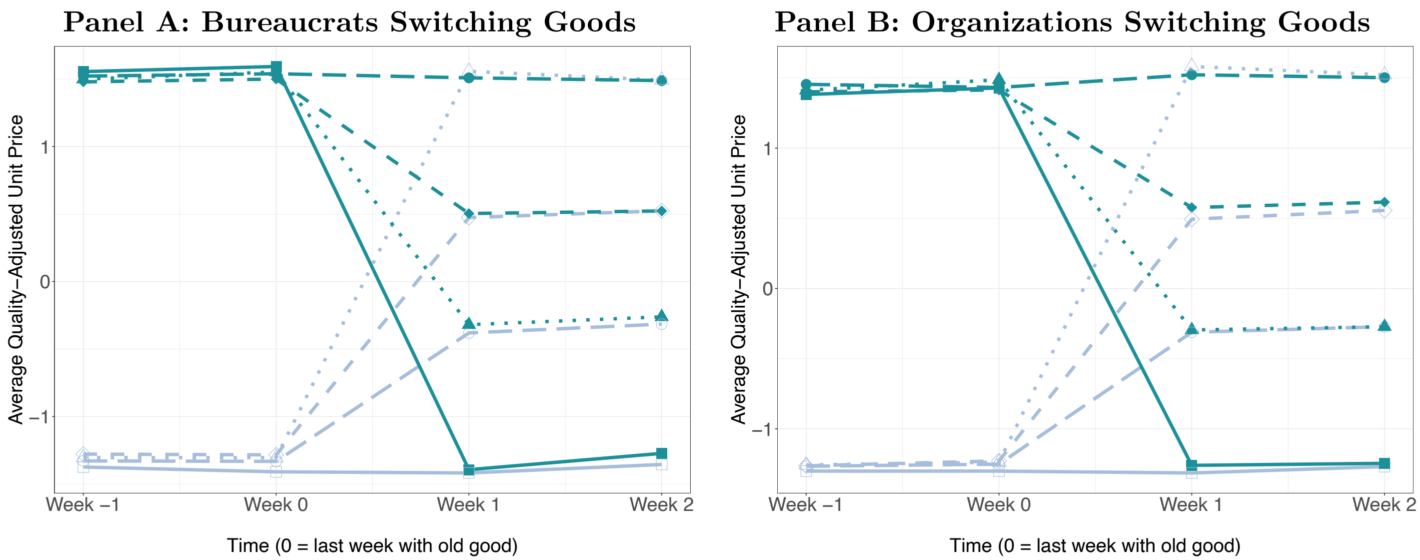
Note: Each panel in the figure is analogous to Figure 1 (see notes to that figure for details of construction), with the following changes. In Panel A, rather than requiring the bureaucrat-organization pair to work together in two separate weeks, we require the pair to work together on two separate days. In Panel B, two separate fortnights; and in Panel C, two separate months. In Panel D we require bureaucrat-organization pairs to work together in three separate weeks.

Figure D.3. Robustness of Event Studies to Design Choices (2)



Note: Each panel in the figure is analogous to Figure 1 (see notes to that figure for details of construction), with the following changes. In Panel A, we categorize bureaucrats by terciles rather than quartiles. In panel B, we construct quartiles by ranking bureaucrats based on the entire sample period rather than each semester separately. Rather than defining spells as weeks separated by fewer than 400 days as in Figure 1, we require them to be separated by 350 days (Panel C) or 450 days (Panel D).

Figure D.4. Event Study of Procurement Prices Around Times Bureaucrats and Organizations Switch Goods



Note: Each panel in the figure is analogous to Figure 1 that studies price changes around the time that organizations switch the bureaucrat making their purchases (see notes to that figure for details of construction). Panel A shows price changes around the time that bureaucrats switch the good they are purchasing. Panel B shows price changes around the time that organizations switch the good they are purchasing.

Table D.1—Event Studies Summary Statistics

			Mean Log Residuals of Bureaucrat Movers				Mean Weeks Betw. Cols:		
Origin/destination Quartile*	Number of Moves (1)	Number of Observations (2)	Week -1 (3)	Week 0 (4)	Week 1 (5)	Week 2 (6)	(3)-(4) (7)	(4)-(5) (8)	(5)-(6) (9)
1 to 1	5,681	231,363	-0.302	-0.369	-0.354	-0.276	12.184	28.487	11.622
1 to 2	5,104	219,228	-0.181	-0.242	-0.101	-0.039	12.201	24.202	12.654
1 to 3	3,317	146,507	-0.197	-0.202	0.050	0.004	13.558	28.907	13.019
1 to 4	1,725	69,720	-0.153	-0.138	0.239	0.246	13.706	36.518	16.572
2 to 1	5,532	221,321	-0.051	-0.088	-0.207	-0.184	13.017	26.492	12.921
2 to 2	8,156	413,847	-0.042	-0.062	-0.024	-0.027	12.029	26.914	12.756
2 to 3	6,121	276,032	-0.021	-0.037	0.087	0.040	12.430	27.855	14.562
2 to 4	2,309	88,520	0.031	0.008	0.254	0.185	12.946	37.727	16.009
3 to 1	3,588	139,430	0.066	0.018	-0.113	-0.166	15.803	25.014	11.276
3 to 2	5,898	259,835	-0.003	0.050	0.018	0.000	13.317	24.493	12.360
3 to 3	5,740	256,476	0.019	0.088	0.125	0.130	15.624	25.736	13.850
3 to 4	2,868	116,791	0.197	0.178	0.290	0.228	13.316	30.498	16.971
4 to 1	1,414	58,123	0.097	0.115	-0.103	-0.065	15.513	31.636	12.163
4 to 2	1,665	73,036	0.098	0.137	-0.001	0.123	15.580	29.804	12.145
4 to 3	2,250	93,164	0.204	0.333	0.248	0.204	15.377	30.858	13.202
4 to 4	2,618	117,693	0.320	0.380	0.391	0.367	15.668	27.312	14.968
Totals	63,986	2,781,086							

Note: The table shows information on events in which organizations switch bureaucrats used in Figure 1. The sample used is the All Products-Analysis Sample summarized in Table 1. Events are defined using the procedure described in detail in Sub-section IV.A. We define an employment spell as a sequence of at least two weeks a bureaucrat-organization pair conducts purchases together, with the weeks less than 400 days apart. Wherever possible, we then match an employment spell (event time $\leq$ 0) with the earliest future spell (event time ¿ 0) involving the same organization but a different bureaucrat. This change of bureaucrats then constitutes an event (event time = 0). We classify the two bureaucrats involved in the event using the average quality-adjusted price they achieve in purchases they make for *other* organizations during the half-year that the spell ends (for the earlier spell) or starts (for the later spell). We run equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. This regression regresses the price achieved in an auction on log quantity, good fixed effects, month fixed effects, interactions between 2-digit HS product categories, years, regions, and lot size, as explained in detail in Sub-section IV.B. Using the price residuals, we then classify bureaucrats by the average they achieve in purchases they make for other organizations. We assign this bureaucrat-average quality-adjusted price to the relevant quartile of the distribution of the average quality-adjusted prices of all bureaucrats that themselves are part of an event in the same half-year as the bureaucrat in question.

Table D.2—Comparing Event Study Data

	Full Sample	Event Study Data
(1) # of Bureaucrats	37,722	6,349
(2) # of Organizations	44,560	17,249
(3) # of Connected Sets	616	289
(4) # of Bureaucrats with >1 Org.	11,320	4,806
(5) # of Organizations with >1 Bur.	37,536	16,969
(6) Mean # of Bureaucrats per Org.	6.02	4.81
(7) Mean # of Organizations per Bur.	7.12	13.1
(8) # of Federal Organizations	1,583	147
(9) # of Regional Organizations	15,530	6,919
(10) # of Municipal Organizations	27,440	10,182
(11) # of Health Organizations	7,231	4,215
(12) # of Education Organizations	25,271	9,273
(13) # of Internal Affairs Organizations	668	136
(14) # of Agr/Environ Organizations	255	92
(15) # of Other Organizations	11,135	3,533
(16) # of Goods	15,649	12,964
(17) Mean # of Goods Per Bur.	93.2	124
(18) # of Regions	86	86
(19) Mean # of Regions per Bur.	1	1
(20) # of Auction Requests	1,871,717	378,539
(21) Mean # of Requests per Bur.	49.6	59.6
(22) Mean # of Applicants	2.94	3.14
(23) Mean # of Bidders	2.07	2.07
(24) Mean Reservation Price	0.291	0.293
(25) Quantity Mean	1,124	1,022
Median	27	35.3
SD	174,951	115,040
(26) Total Price Mean (bil. USD)	81.2	64
Median	4.74	3.3
SD	482	422
(27) Unit Price Mean (bil. USD)	55.6	38.4
Median	0.18	0.086
SD	19,168	1,214
(28) Mean # of Contract Renegotiations (log)	0.133	0.117
(29) Mean Size of Cost Over-run	-0.002	-0.002
(30) Mean Length of Delay in Days (log)	0.057	0.082
(31) Mean 1[End User Complained about Contract]	0.001	0.001
(32) Mean 1[Contract Cancelled]	0.009	0.016
(33) Mean 1[Product is of Substandard Quality]	0.009	0.006
(34) # of Observations	16,348,331	4,042,533
(35) Total Procurement Volume (bil. USD)	629	122

Note: The table reports summary statistics for two samples. 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.

E: Additional Results on Variance Decomposition

This appendix presents additional results on the variance decomposition discussed in sections IV.B and IV.C. Sub-section E.E1 presents further evidence in support of the log-linear specification (3) used in the variance decomposition. Sub-section E.E2 presents additional results showing the robustness of the findings to various design choices.

E1 Misspecification

The model we have estimated assumes that the price achieved is approximately log-linear in the bureaucrat and organization effects. Three pieces of evidence suggest that match-based forms of endogenous mobility that would violate the identifying assumptions underlying our interpretation of the results from our empirical model rarely occur in Russian public procurement. First, the event studies in Sub-section IV.A provide direct visual evidence that the price paid is approximately log-linear in the bureaucrat and organization effects. We saw no evidence of sorting on match effects in Figure 1.

Second, a direct piece of evidence in support of the log-linearity assumption comes from studying the distribution of the residuals across bureaucrat and organization effect deciles. If the log-linear specification was substantially incorrect, we would expect to see systematic patterns in the residuals. For example, positive match effects would lead the residuals to be large when the bureaucrat and organization are both in the top deciles of effectiveness. Panel A of Figure E.1 shows a heat map of residuals. The map reveals no clear patterns in the residuals. Panel B shows an analogous heat map of residuals from running (3) in levels rather than logs. The figure provides clear evidence that such a model is mis-specified, leading to systematically large residuals especially in the top right of the figure, where both the bureaucrat and organization are in the top deciles of effectiveness.

Third, we reestimate equation (3) but include fixed effects for each bureaucrat-organization pair, allowing for arbitrary patterns of complementarity between bureaucrats and organizations (see also Card, Heining and Kline, 2013). If there are indeed strong or moderate match effects that our model omits, then we expect this pair effect model to fit significantly better. The pair effect model does not fit the data much better than our baseline model: adding pair effects decreases the RMSE of the residuals from 1.147 to 1.121 and increases the adjusted R^2 from 0.963 to 0.964, and the pair effects have a much smaller variance than the procurer effects from the log-linear model (results available from the authors upon request).

Overall, we do not find evidence supporting a rejection of our log-linearity assumption.

E2 Robustness to design choices

In figure 2 discussed in section IV.D, we argue that our results are robust to focusing on more homogeneous subsets of goods in our sample. We use the measure of the scope for quality differentiation developed by Sutton (1998). As an alternative, we repeat the exercise using the measure developed by Khandelwal (2010) in figure E.2. The results are extremely similar. In particular, the share of the variation in prices explained by the bureaucrats and organizations remains constant as we increase the the degree of good homogeneity moving from right to left.

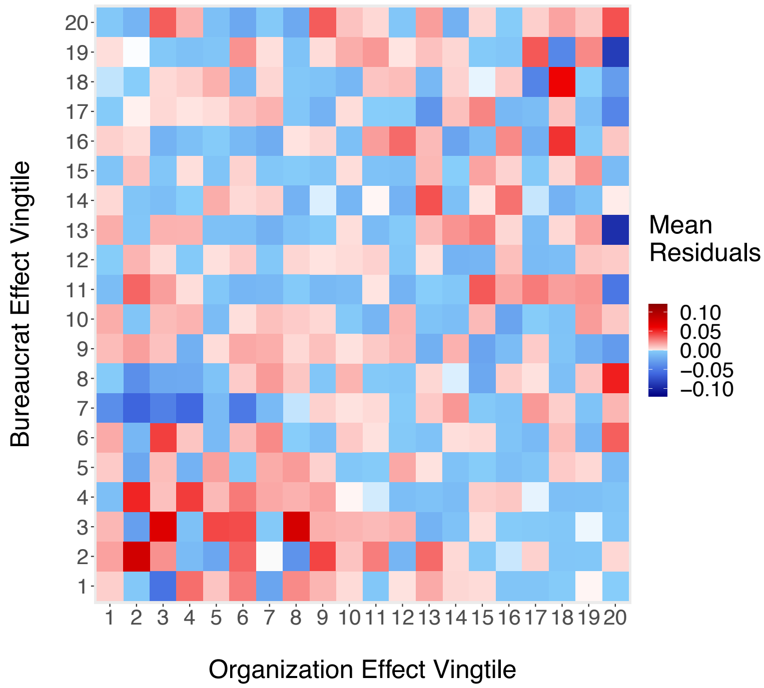
As we discuss in section IV.D, prices are the most important outcome in procurement, but not the only one, and so we also study the impact of bureaucrats and organizations on the spending quality measures described in section II.B. We argue that these outcomes are endogenous to the bureaucrats and organizations in charge of procurement, and hence do not belong as controls in the variance decomposition. Nevertheless, in Appendix Table E.1 we re-estimate the variance decomposition including the spending quality outcomes as controls, and show that the results are essentially unchanged from our baseline specification in table 2 (for example, the standard deviation of the joint effect of the buyers goes from 0.499 down only to 0.484).

As discussed in section IV.B, bureaucrat- and organization- effects can only be estimated within sets of organizations connected by bureaucrats switching between them — connected sets. In our main analysis we pool the connected sets. As a robustness check, here we present results using only the largest connected set in the data. Table E.2 presents summary statistics of this largest connected set. The sample is broadly comparable to the main sample. Table E.3 shows the results of the variance decomposition in the largest connected set. The results are very similar to the main sample. The fixed effects, split-sample and shrinkage methods all attribute roughly the same share of the variation to the bureaucrats and organizations as in the full sample. The covariance shrinkage method attributes a bit less, 30%, slightly less than in the full sample. This gives us confidence that our results apply well beyond the lasrgest connected set.

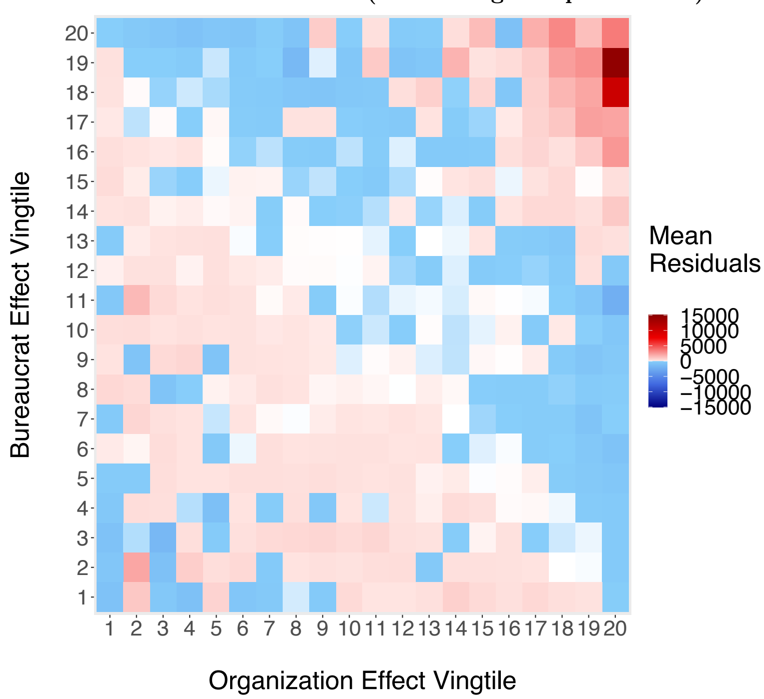
Section IV.B and appendix A describe the steps we took to build our analysis sample. Table E.4 shows the robustness of our estimates to the three main design choices. Column (1) replicates the findings in column (1) of table 2. Columns (2) and (3) use lower (45th percentile) and higher (55th percentile) thresholds of confidence to identify correctly classified items, respectively. Columns (4) and (5) trim fewer (top and bottom 2.5%) and more (top and bottom 10%) outlier observations for each good. Column (6) uses the Support Vector Machine classifier and column (7) uses the hierarchical classifier. All details are described in section A. As the table reveals, the results are remarkably stable across samples, reassuring us that our results are not driven by our sample building strategy.

Figure E.1. Correlation of Residuals with Estimated Bureaucrat and Organization Effects

Panel A: Prices in Logs (Main Specification)

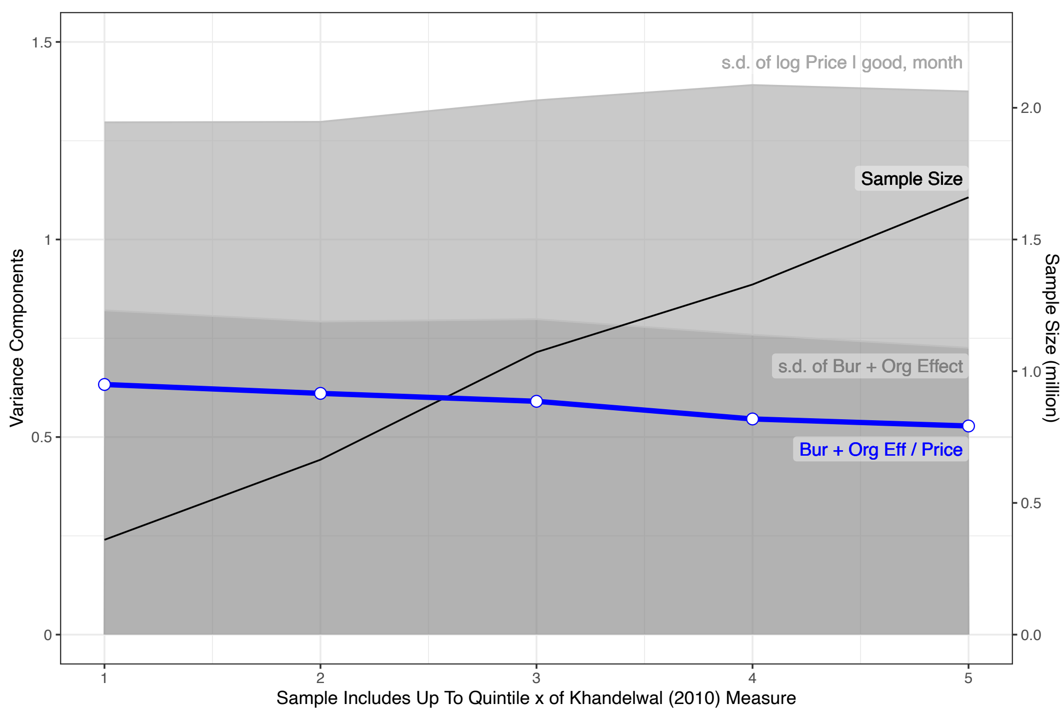


Panel B: Prices in Levels (Illustrating Misspecification)



Note: The figure presents heatmaps of averages of the residuals from the estimation of equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ — in logs (Panel A) and in levels (Panel B). The residuals are binned by vingtiles of the estimated bureaucrat effect $\hat{\alpha}_b$ and organization effect $\hat{\psi}_j$ within each connected set. The sample used is the Analysis Sample (All Products) summarized in Table 1.

Figure E.2. Robustness to Using Subsamples of Increasingly Heterogeneous Goods (Khandelwal (2010) Measure)



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 figure uses the sub-set of the sample that we can match to the scope-for-quality-differentiation ladder developed by Khandelwal (2010). 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.

Table E.1—Robustness of Variance Decomposition of Prices to Including Spending Quality Controls

	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)	1.196	(0.0293)	1.259	(0.0313)	0.820	0.425
(2) s.d. of Organization Effects (across orgs)	1.126	(0.0398)	1.188	(0.0472)	0.778	0.371
(3) s.d. of Bureaucrat Effects (across items)	0.787	(0.0304)	0.836	(0.0436)	0.591	0.266
(4) s.d. of Organization Effects (across items)	0.922	(0.0431)	0.984	(0.0561)	0.700	0.342
(5) Bur-Org Effect Correlation (across items)	-0.721	(0.0176)	-0.565	(0.0395)	-0.664	0.299
(6) s.d. of Bur + Org Effects Within CS (across items)	0.651	(0.0165)	0.662	(0.0202)	0.538	0.492
(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 controls sets estimated by implementing the variance decomposition in equation (4) extended to include our spending quality measures as controls. 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): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. 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 E.2—Summary Statistics - Largest Connected Set

	All Products		Pharmaceuticals Subsample	
	(1) No Preferences Analysis Sample	(2) Largest Connected Set	(3) No Preferences Analysis Sample	(4) Largest Connected Set
(1) # of Bureaucrats	37,722	6,083	2,473	3,087
(2) # of Organizations	44,560	9,001	1,866	1,899
(3) # of Connected Sets	616	1	129	1
(4) # of Bureaucrats with >1 Org.	11,063	1,792	926	31
(5) # of Organizations with >1 Bur.	37,306	7,213	1,449	637
(6) Mean # of Bureaucrats per Org.	5.59	6.42	4.32	1.65
(7) Mean # of Organizations per Bur.	6.6	9.51	3.26	1.02
(8) # of Federal Organizations	1,583	166	26	478
(9) # of Regional Organizations	15,530	3,513	1,599	1,270
(10) # of Municipal Organizations	27,440	5,320	241	151
(11) # of Health Organizations	7,231	1,604	1,705	1,579
(12) # of Education Organizations	25,271	4,892	61	36
(13) # of Internal Affairs Organizations	668	98	3	102
(14) # of Agr/Environ Organizations	255	59	1	25
(15) # of Other Organizations	11,135	2,348	96	157
(16) # of Goods	14,875	12,048	3,861	3,713
(17) Mean # of Goods Per Bur.	72.5	83.8	42.5	22.8
(18) # of Regions	86	28	85	85
(19) Mean # of Regions per Bur.	1	1	1	1
(20) # of Auction Requests	1,199,363	248,999	42,875	19,817
(21) Mean # of Requests per Bur.	31.8	40.9	17.3	6.42
(22) Mean # of Applicants	3.04	3.11	2.65	2.39
(23) Mean # of Bidders	1	1	1.94	1.88
(24) Mean Reservation Price	0.291	0.288	0.303	0.304
(25) Quantity Mean	1,053	951	1,719	333
Median	25	30	45	35
SD	90,917	40,257	172,144	2,972
(26) Total Price Mean (bil. USD)	80.1	70.8	91.1	189
Median	4.32	3.72	6.7	5.69
SD	493	460	493	259
(27) Unit Price Mean (bil. USD)	61.3	48.8	25.4	11.6
Median	0.167	0.132	0.18	0.169
SD	23,015	2,076	265	138
(28) Mean # of Contract Renegotiations (log)	0.121	0.12	0.142	0.168
(29) Mean Size of Cost Over-run	-0.002	-0.001	-0.003	-0.004
(30) Mean Length of Delay in Days (log)	0.064	0.07	0.076	0.078
(31) Mean 1[End User Complained about Contract]	0.001	0.001	0	0.001
(32) Mean 1[Contract Cancelled]	0.012	0.013	0.016	0.015
(33) Mean 1[Product is of Substandard Quality]	0.005	0.003	0.058	0.112
(34) # of Observations	11,339,187	2,258,081	181,963	108,376
(35) Total Procurement Volume (bil. USD)	395	54.1	9.38	5.14

Note: The table reports summary statistics for four 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. Analysis Sample denotes all unpreferenced auctions in connected sets that fulfill three restrictions: singleton bureaucrat-organization, bureaucrat-good, and organization-good pairs are removed; each procurer (bureaucrats and organizations) implements a minimum of five purchases; and connected sets have at least three bureaucrats and organizations. Largest Connected Set is the largest connected set from the Analysis Sample (as measured by the number of organizations). 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 E.3—Share of Variance of Procurement Prices explained by Bureaucrats and Organizations: Largest Connected Set

	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)	1.007	(0.0327)	1.051	(0.041)	0.698	0.388
(2) s.d. of Organization Effects (across orgs)	1.184	(0.0904)	1.231	(0.112)	0.849	0.329
(3) s.d. of Bureaucrat Effects (across items)	0.580	(0.0384)	0.618	(0.0469)	0.453	0.242
(4) s.d. of Organization Effects (across items)	0.841	(0.0956)	0.906	(0.118)	0.681	0.211
(5) Bur-Org Effect Correlation (across items)	-0.688	(0.0285)	-0.486	(0.0786)	-0.658	0.467
(6) s.d. of Bur + Org Effects Within CS (across items)	0.610	(0.0442)	0.631	(0.0524)	0.513	0.388
(7) s.d. of log unit price	2.165		2.165		2.165	2.165
(8) s.d. of log unit price | good, month	1.207		1.207		1.207	1.207
(9) Adjusted R-squared	0.964		0.964		0.964	0.964
(10) Number of Bureaucrats	6,083		6,083		6,083	6,083
(11) Number of Organizations	9,001		9,001		9,001	9,001
(12) Number of Bureaucrat-Organization Pairs	57,822		57,822		57,822	57,822
(13) Number of Connected Sets	1		1		1	1
(14) Number of Observations	2,258,081		2,258,081		2,258,081	2,258,081

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-Largest Connected Set Sample summarized in Table E.2. 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): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. 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 E.4—Variance Decomposition Results: Robustness to Sample Definition

Machine Learning Method	LR	LR	LR	LR	LR	SVM	HM
Classification Confidence Threshold	50	45	55	50	50	50	50
Outlier Trimming	5	5	5	2.5	10	5	5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) s.d. of Bureaucrat Effects (across burs)	1.197	1.255	1.207	1.602	0.914	1.274	1.182
(2) s.d. of Organization Effects (across orgs)	1.130	1.190	1.147	1.590	0.846	1.223	1.108
(3) s.d. of Bureaucrat Effects (across items)	0.788	0.855	0.826	1.234	0.592	0.890	0.778
(4) s.d. of Organization Effects (across items)	0.927	0.989	0.945	1.438	0.736	1.003	0.904
(5) Bur-Org Effect Correlation (across items)	-0.720	-0.746	-0.738	-0.761	-0.717	-0.753	-0.713
(6) s.d. of Bur + Org Effects Within CS (across items)	0.655	0.669	0.650	0.943	0.517	0.673	0.648
(7) s.d. of log unit price	2.188	2.203	2.188	2.417	1.854	2.194	2.188
(8) s.d. of log unit price | good, month	1.280	1.299	1.280	1.411	1.094	1.250	1.282
(9) Adjusted R-squared	0.963	0.962	0.963	0.958	0.970	0.965	0.963
(10) Number of Bureaucrats	37,722	38,154	37,722	40,959	34,438	37,893	37,563
(11) Number of Organizations	44,560	44,736	44,560	46,751	41,895	44,759	44,506
(12) Number of Bureaucrat-Organization Pairs	248,898	250,364	248,898	265,414	231,752	250,394	248,787
(13) Number of Connected Sets	616	614	616	619	606	618	618
(14) Number of Observations	11,339,187	11,362,565	11,339,187	12,088,012	10,016,651	11,365,756	11,343,315

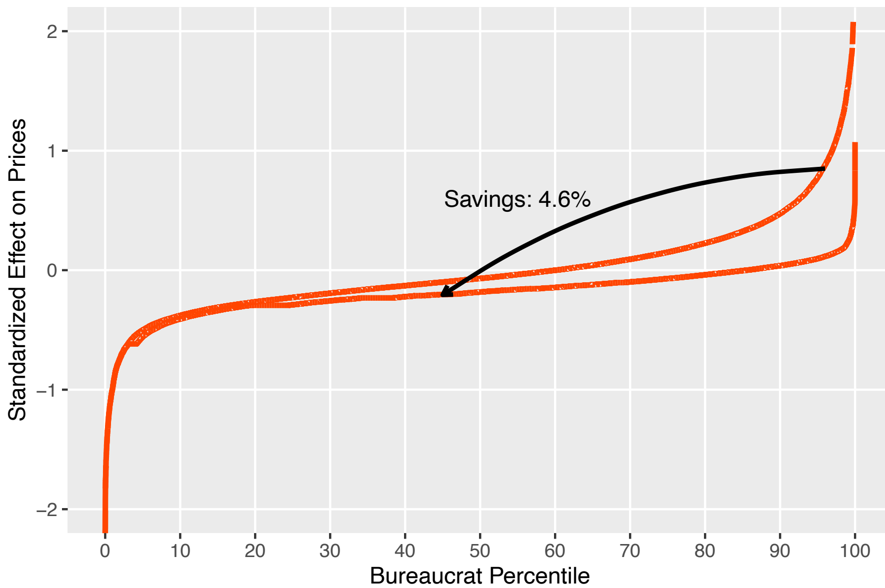
Note: The table shows the components of the variance due to bureaucrats, organizations, and connected sets estimated by implementing the variance decomposition in equation (4) in different samples. The decomposition uses the fixed effect estimates from equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta}+\alpha_{b(i,j)}+\psi_j+\gamma_{s(b,j)}+\varepsilon_i$. Column (1) replicates the findings in column (1) of table 2. Columns (2) and (3) use lower (45th percentile) and higher (55th percentile) thresholds of confidence to identify correctly classified items, respectively. Columns (4) and (5) trim fewer (top and bottom 2.5%) and more (top and bottom 10%) outlier observations for each good. Column (6) uses the Support Vector Machine classifier and column (7) uses the hierarchical classifier. All details are described in Appendix A.

E3 Crude Counterfactuals

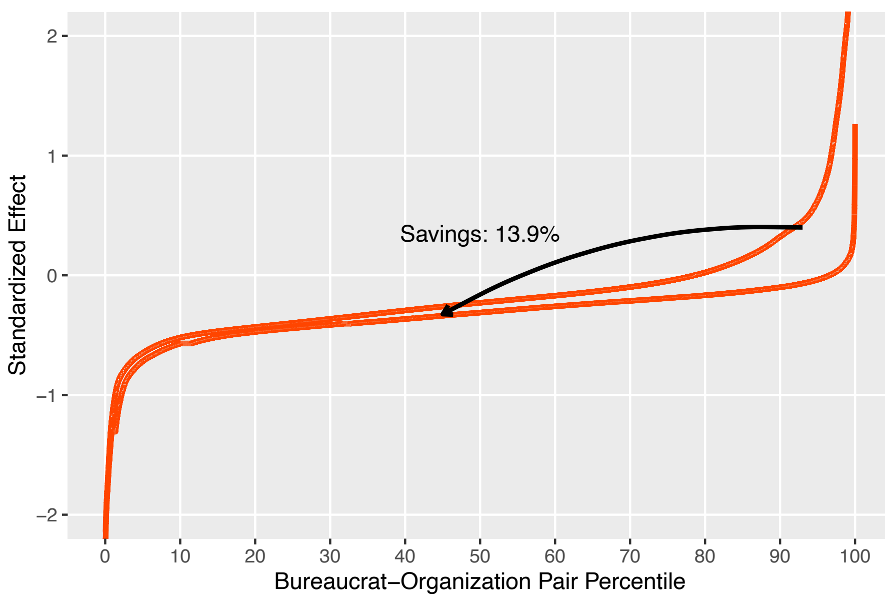
Our large estimates of the share of variation in performance attributable to bureaucrats and organizations have correspondingly dramatic implications for the scope of potential savings from improving the effectiveness of the bureaucracy. To illustrate the magnitude, we can consider simple counterfactual bureaucracies in which bureaucrats and/or organizations with low effectiveness are improved, for example through changes in recruiting, training of existing bureaucrats, or improved organizational management. Figure E.3 shows two such counterfactuals. Panel A shows the shift in the distribution of bureaucrat effects that would occur if the lowest quartile of bureaucrats were able to be improved to the 75th percentile. This would save the Russian government 4.6 percent of annual procurement expenses. In Panel B we consider moving all bureaucrats *and* organizations below 25th percentile-effectiveness to 75th percentile-effectiveness. The panel shows the distribution of pair (bureaucrat plus organization) effects that would result. The government would save 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.

Figure E.3. Crude Counterfactuals

Panel A: Least Effective 25% of Bureaucrats to 75th Percentile



Panel B: Least Effective 25% of Bureaucrats & Organizations to 75th Percentile



Note: The figure shows the impact of two counterfactual scenarios on the distribution of our estimated price effects. Panel A considers moving all bureaucrats above the 75th percentile of their connected set's distribution of covariance shrunken price effects down to their connected set's 25th percentile. The dashed line shows the distribution of our covariance shrunken estimates of the bureaucrat effects, while the solid line shows the distribution that would result from implementing the counterfactual. Panel B considers moving both all bureaucrats and all organizations above the 75th percentile of their connected set's distribution of covariance shrunken price effects down to their connected set's 25th percentile. The dashed line shows the distribution of bureaucrat-organization pair effects we estimate, while the solid line shows the distribution that would occur in the counterfactual scenario. Overlaid on both panels are the implied aggregate savings.

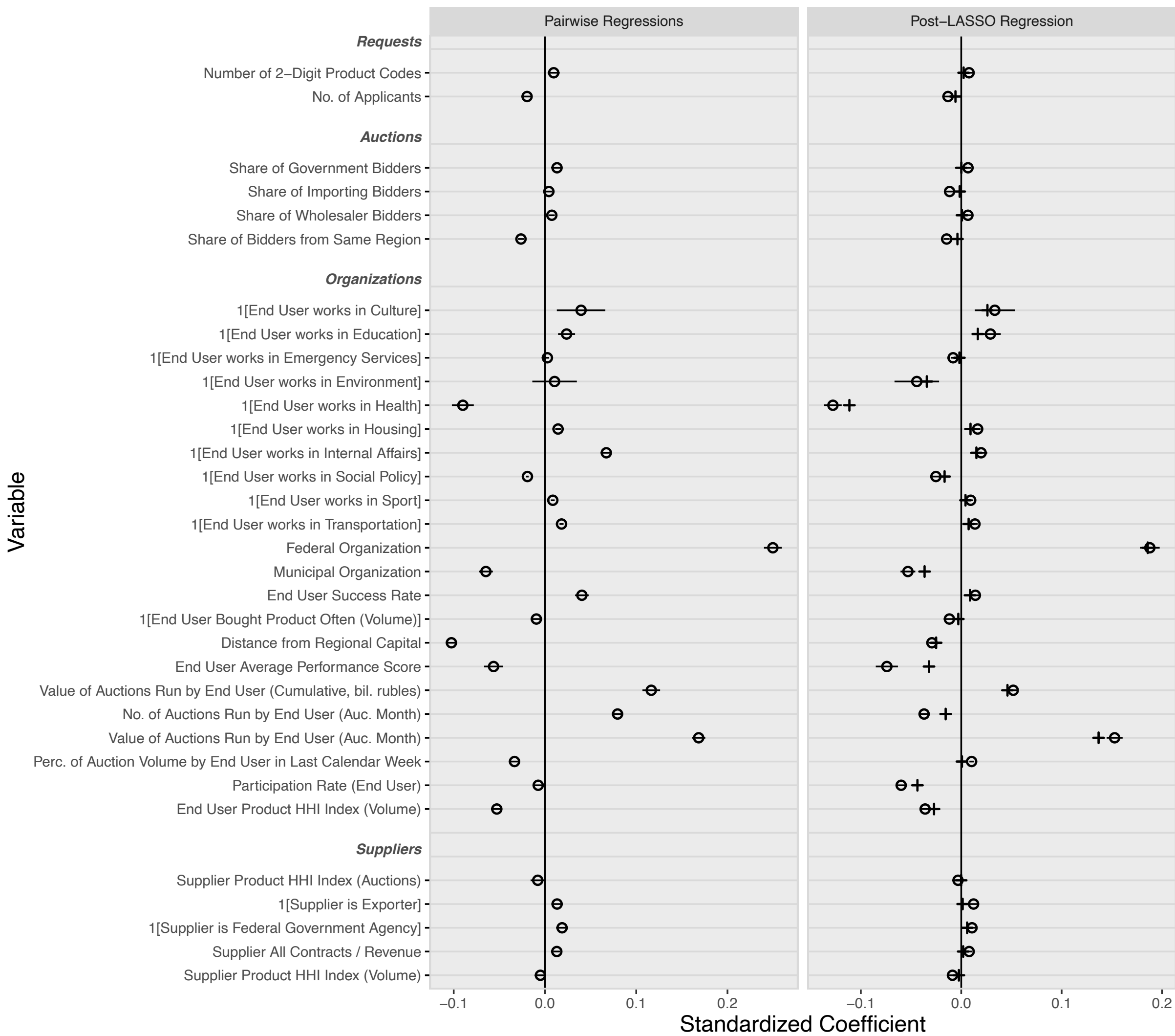
F: Additional Results on What Effective Bureaucracies do Differently

This appendix presents additional results on the variance decomposition discussed in section IV.E. In section IV.E we exploit the richness of our data to analyze the correlates of bureaucratic and organizational effectiveness. To avoid overfitting and for the sake of parsimony, we use a LASSO procedure to first select 30 predictor variables.[81] We then regress each purchase's covariance-shrunk bureaucrat/organization effect on these variables, the purchase's organization effect, and the controls in (3). In Figures 4, 5, and F.1–F.6, the left panels show regression coefficients from a series of bivariate regressions of the bureaucrat/organization price/spending quality effect on each of the selected observables. The right panels show the LASSO coefficients (as crosses) and the coefficients from the multivariate regression of the procurer effects 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 measure of procurer behavior and the causal impact of the procurer.

In the main paper, we present results on correlates of bureaucrats' price (Figure 4) and spending quality (Figure 5) effects. Figures (F.1) and (F.2) present the analogs for organizations. For parsimony we selected 30 predictor variables, but Figures (F.3) – (F.6) extend Figures (4), (5), (F.1) and (F.2) to pick 60 variables instead of 30. 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. Finally, Table F.1 summarizes the data used in this exercise.

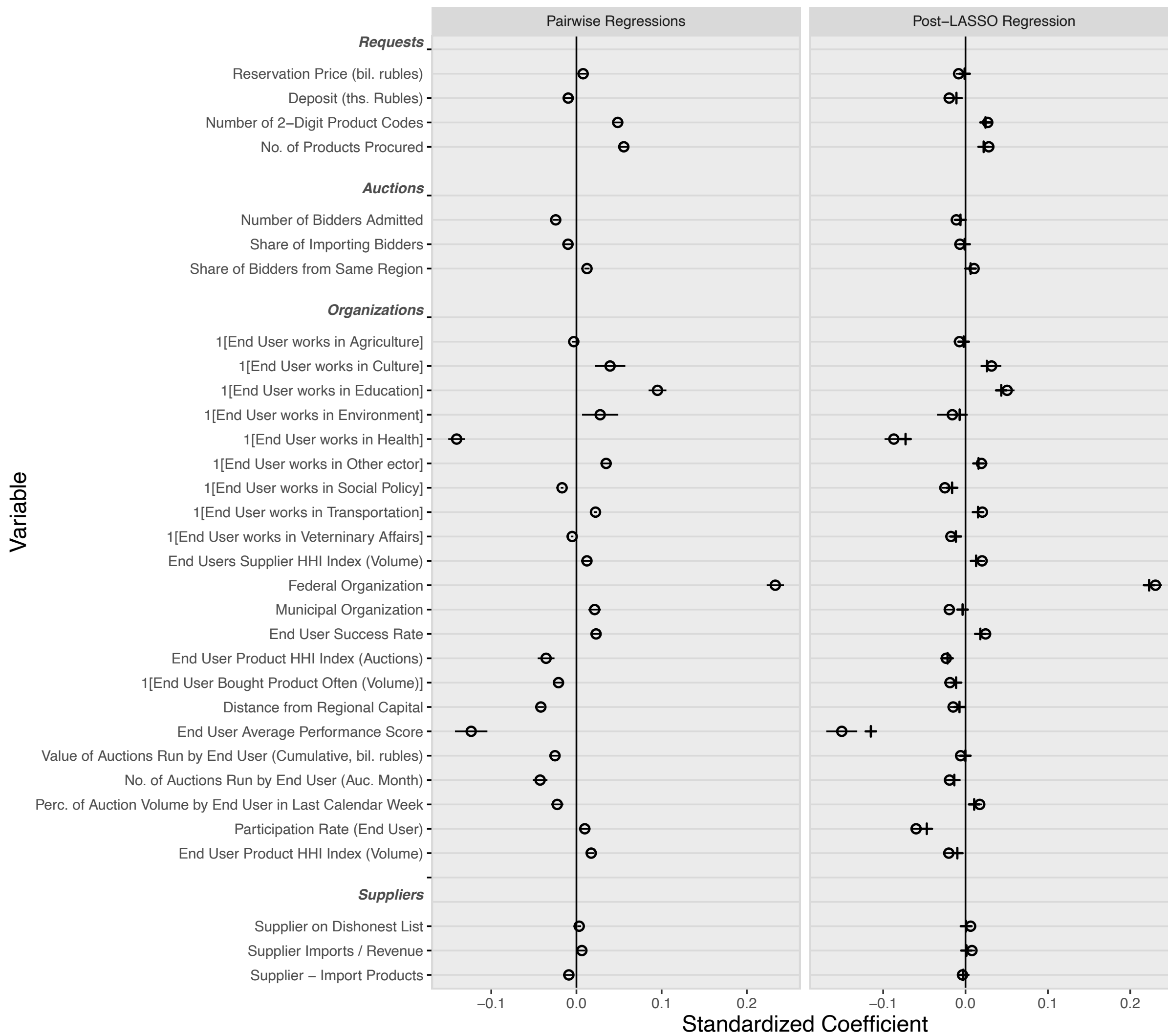
[81]The procedure selects the smallest model with *at least* 30 variables so the actual number varies slightly from figure to figure. Table F.1 shows pairwise coefficients from regressing price-effectiveness on each of the 160 potential explanatory variables we start out with. Tables F.3 and F.4 instead 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 below.

Figure F.1. Correlates of Organization Effectiveness (Price, 30 variables)



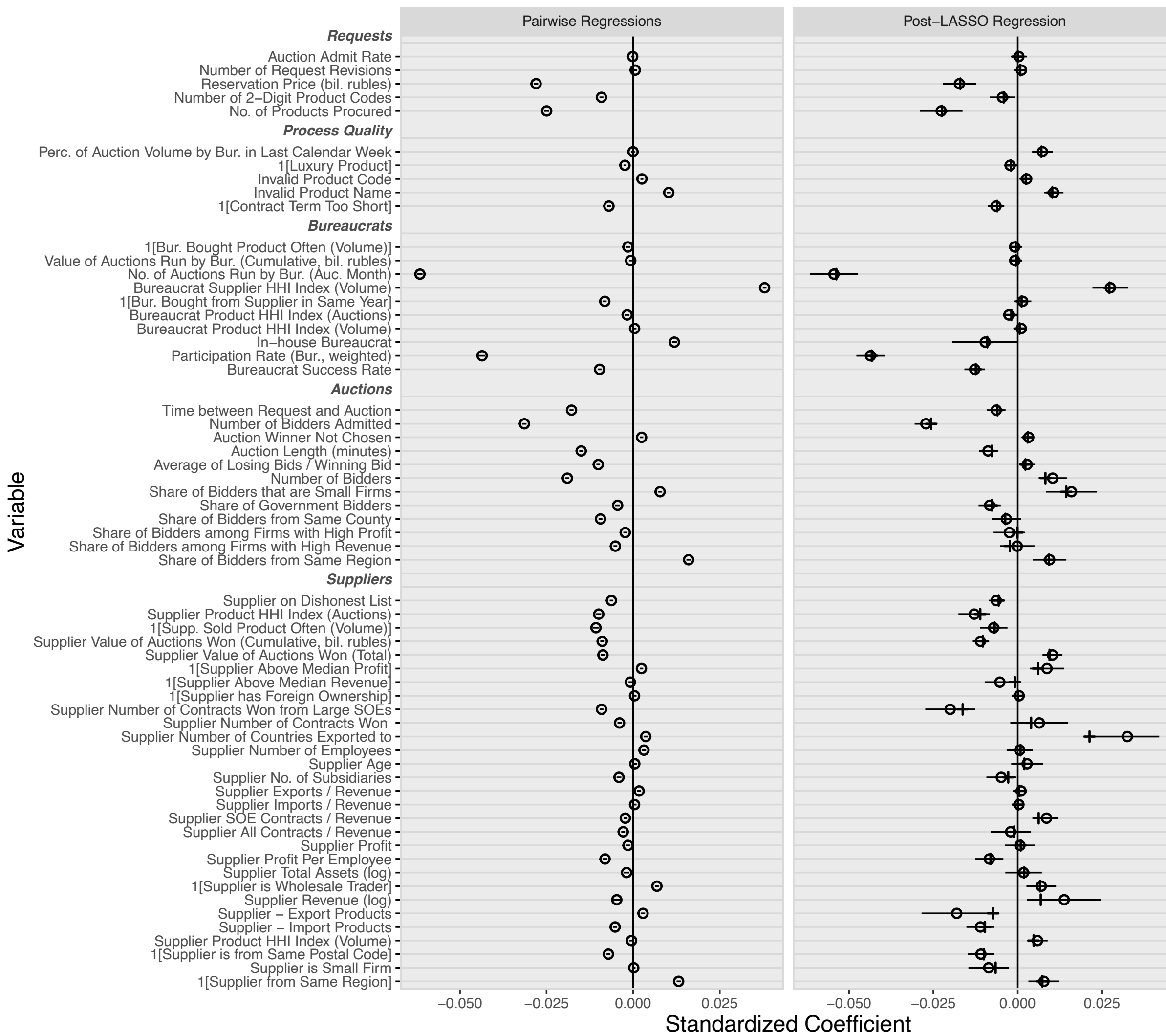
Note: The figure shows the results of regressions of estimated bureaucrat effects $\hat{\alpha}_b$ from estimation of equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ with prices as the outcome on observable characteristics of the purchase procedure followed. 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 from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients from the multivariate regression of the effects on all of the selected variables. All variables are standardized to have unit standard deviation.

Figure F.2. Correlates of Organization Effectiveness (Quality, 30 variables)



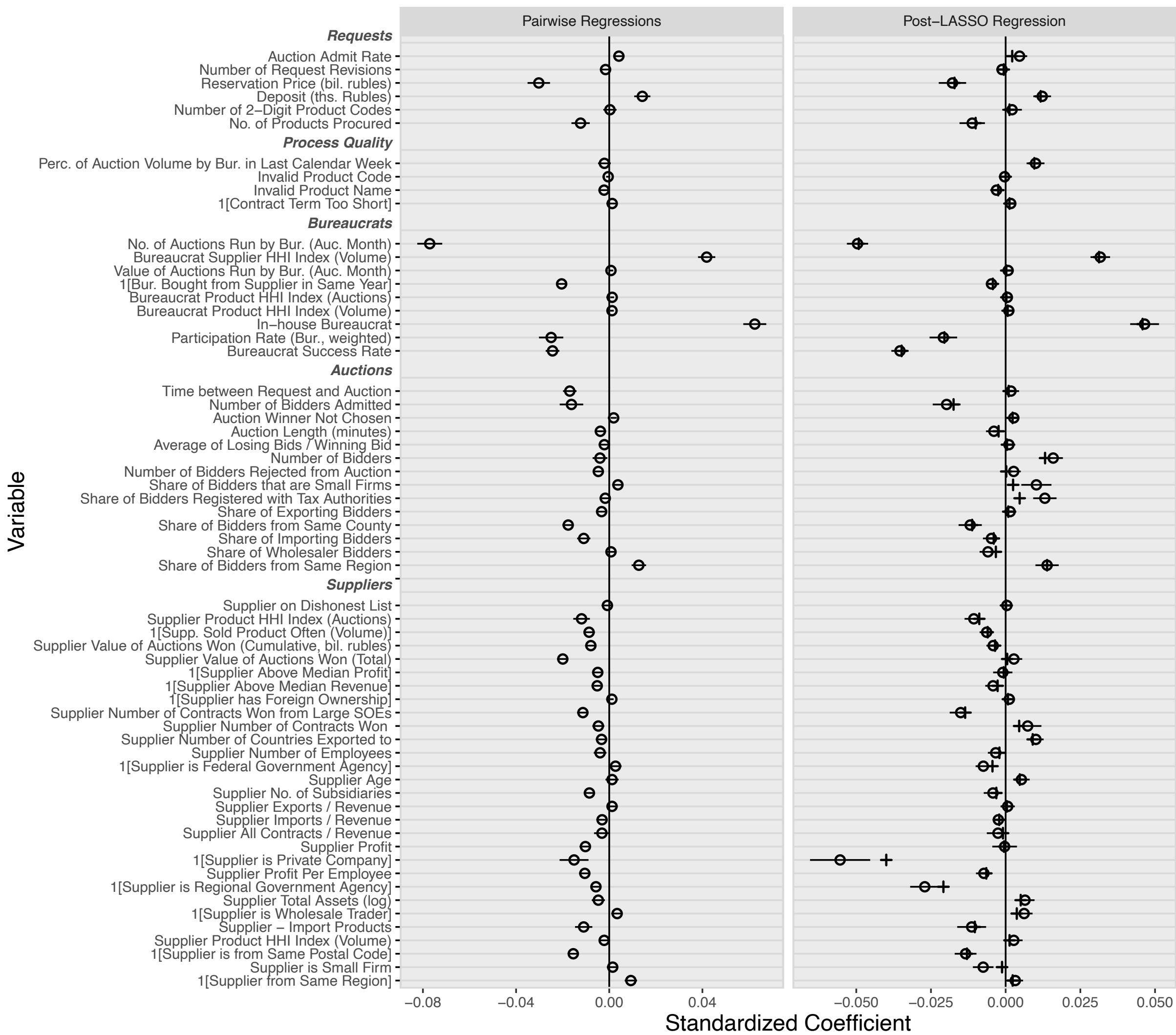
Note: The figure shows the results of regressions of estimated organization effects $\hat{\psi}_j$ from estimation of equation (3): $q_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ with spending quality as the outcome on observable characteristics of the purchase procedure followed. 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 from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients from the multivariate regression of the effects on all of the selected variables. All variables are standardized to have unit standard deviation.

Figure F.3. Correlates of Bureaucrat Effectiveness (Price, 60 variables)



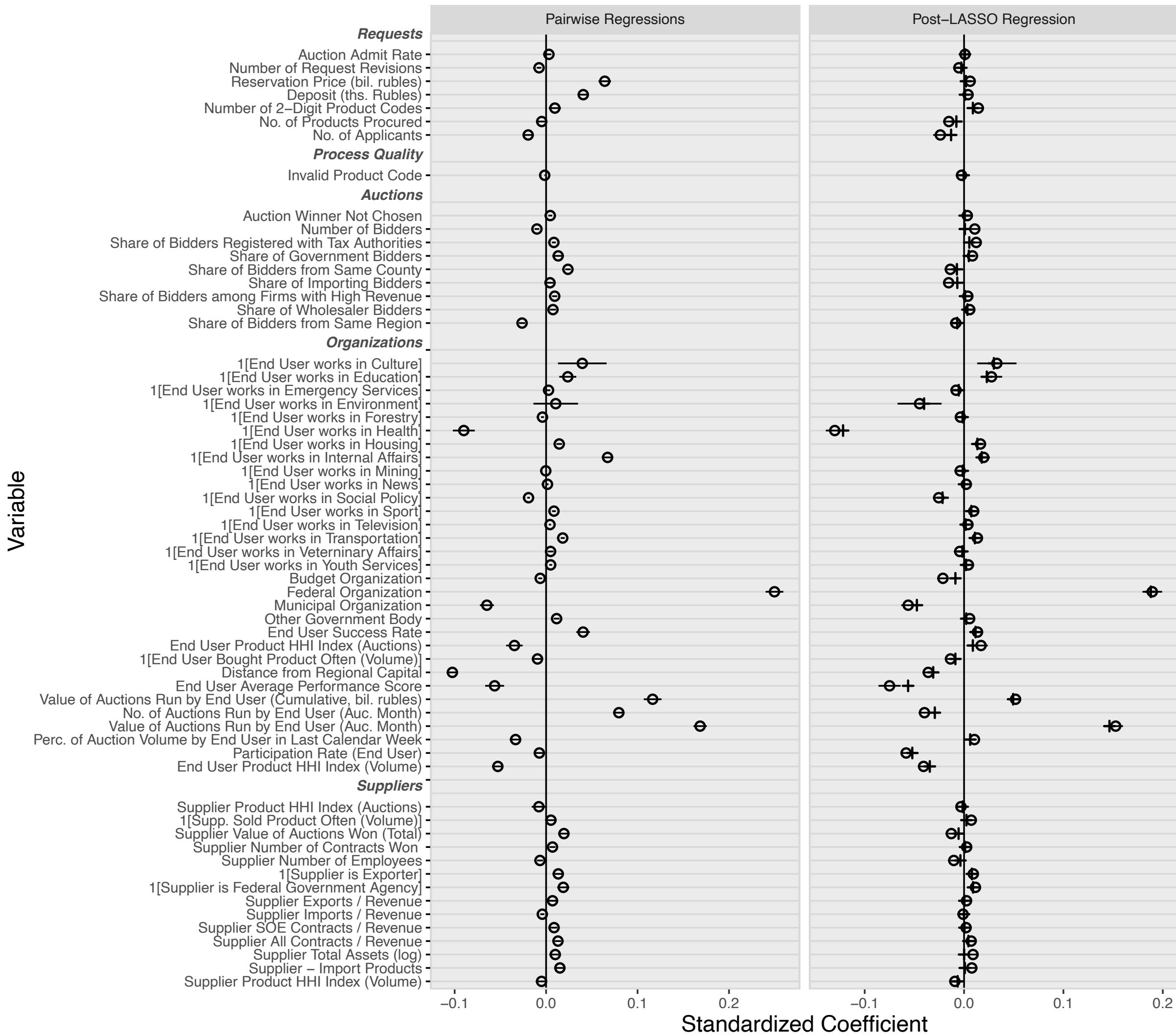
Note: The figure shows the results of regressions of estimated bureaucrat effects $\hat{\alpha}_b$ from estimation of equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta}+\alpha_{b(i,j)}+\psi_j+\gamma_{s(b,j)}+\varepsilon_i$ with price as the outcome on observable characteristics of the purchase procedure followed. We use a LASSO procedure to select 60 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 from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients from the multivariate regression of the effects on all of the selected variables. All variables are standardized to have unit standard deviation.

Figure F.4. Correlates of Bureaucrat Effectiveness (Quality, 60 variables)



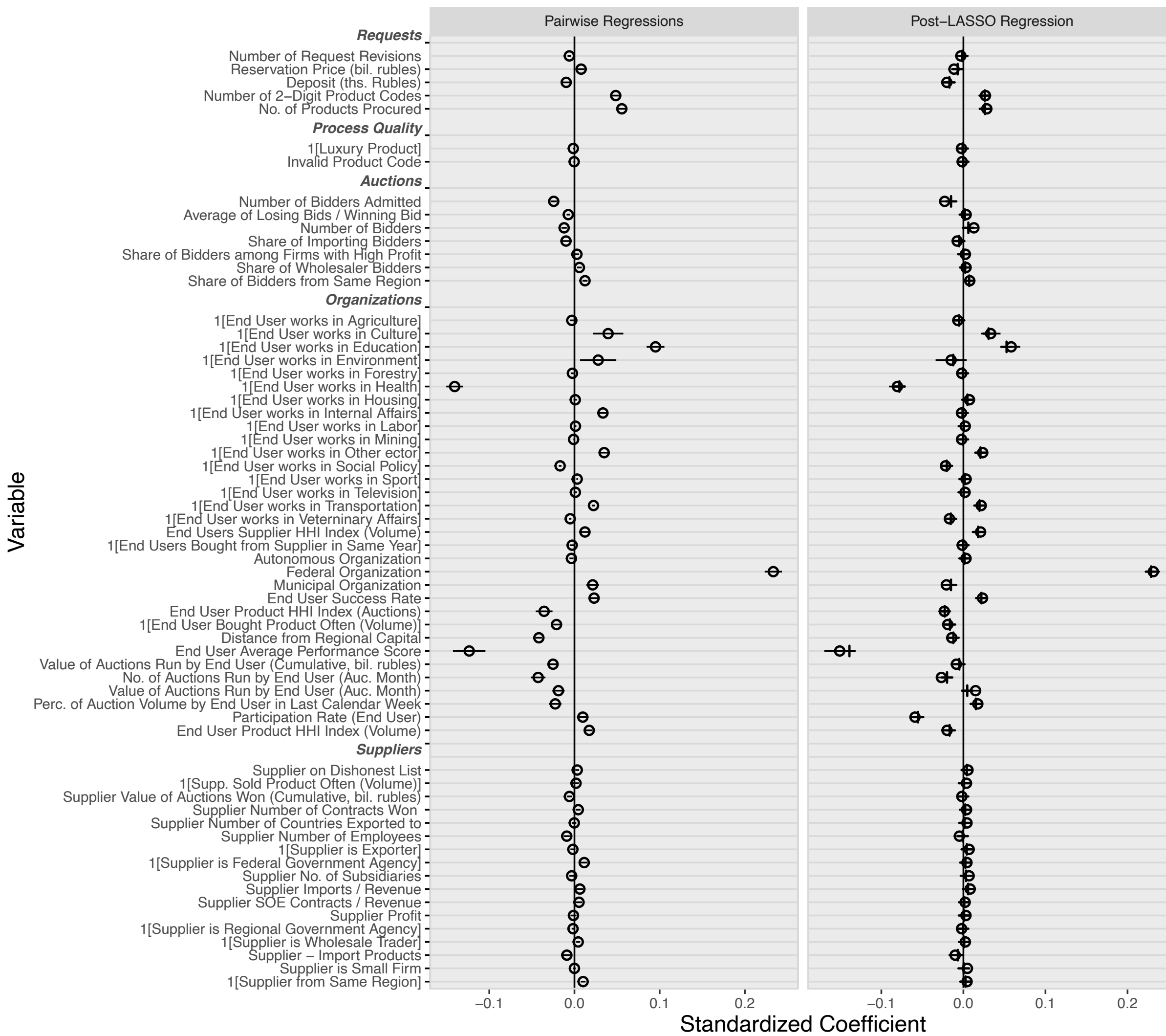
Note: The figure shows the results of regressions of estimated bureaucrat effects $\hat{\alpha}_b$ from estimation of equation (3): $q_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ with spending quality as the outcome on observable characteristics of the purchase procedure followed. We use a LASSO procedure to select 60 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 from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients from the multivariate regression of the effects on all of the selected variables. All variables are standardized to have unit standard deviation.

Figure F.5. Correlates of Organization Effectiveness (Price, 60 variables)



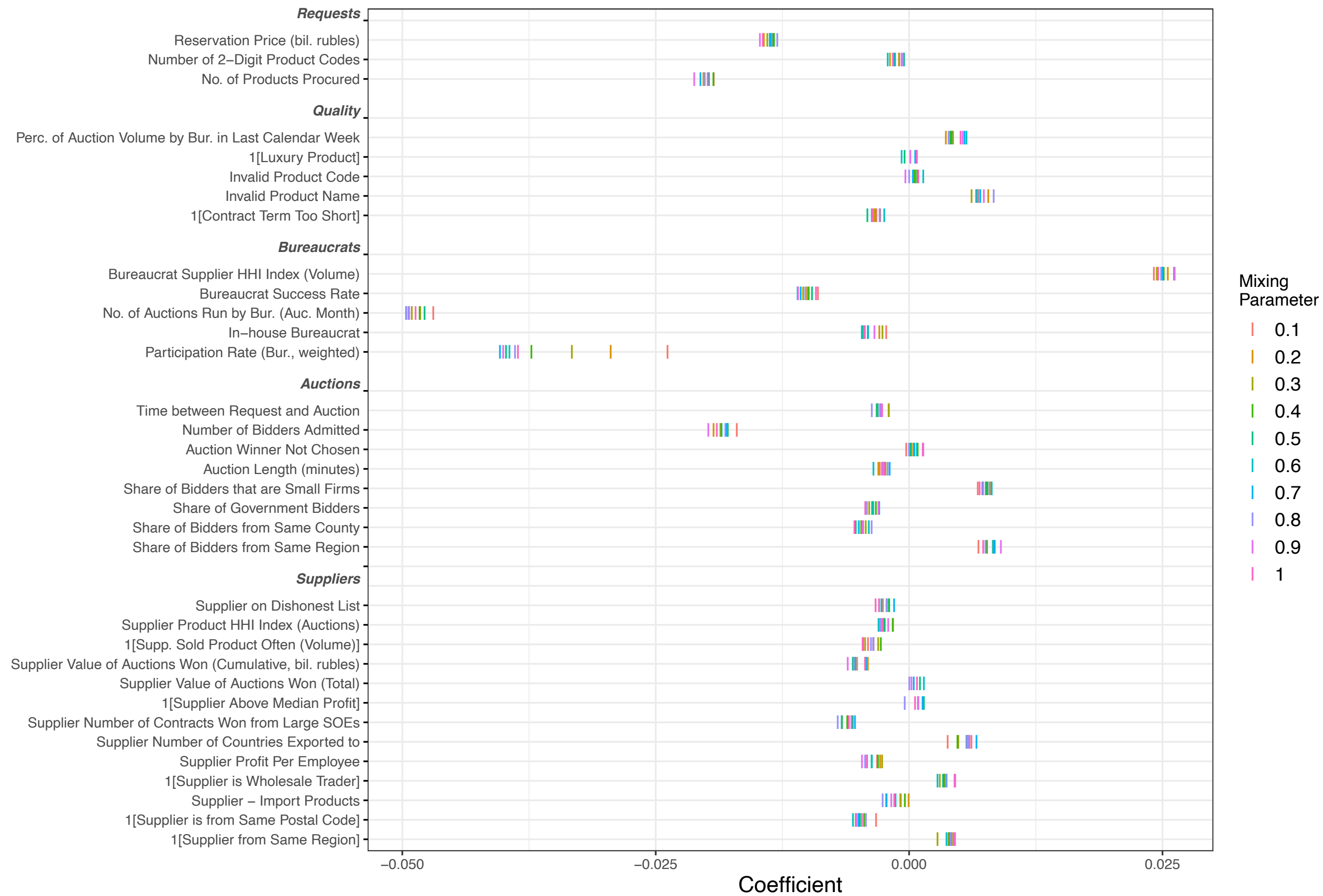
Note: The figure shows the results of regressions of estimated organization effects $\hat{\psi}_j$ from estimation of equation (3): $p_i = \mathbf{X}_i\boldsymbol{\beta}+\alpha_{b(i,j)}+\psi_j+\gamma_{s(b,j)}+\varepsilon_i$ with price as the outcome on observable characteristics of the purchase procedure followed. 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 from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients from the multivariate regression of the effects on all of the selected variables. All variables are standardized to have unit standard deviation.

Figure F.6. Correlates of Organization Effectiveness (Quality, 60 variables)



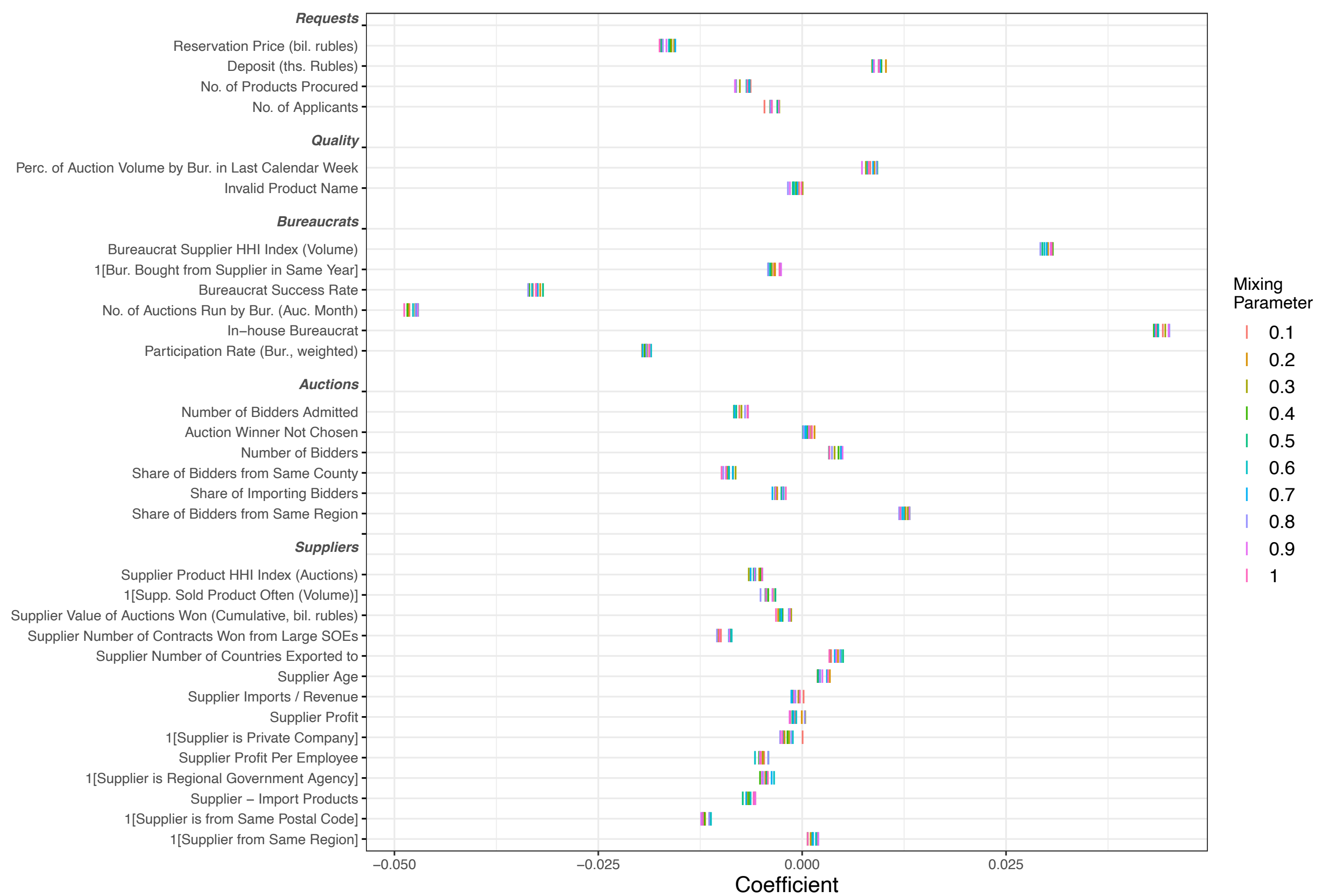
Note: The figure shows the results of regressions of estimated organization effects $\hat{\psi}_j$ from estimation of equation (3): $q_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$ with spending quality as the outcome on observable characteristics of the purchase procedure followed. 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 from a series of bivariate regressions of the bureaucrat effect on each of the selected observables. The right panels show the coefficients from the multivariate regression of the effects on all of the selected variables. All variables are standardized to have unit standard deviation.

Figure F.7. Correlates of Bureaucrat Effectiveness (Price): Elastic Net Regularization Coefficients Across Different Mixing Parameters



Note: The figure shows the coefficients from the elastic net regularization procedure on the estimated bureaucrat effects across different values of the mixing parameters. Each coefficient is represented by a small vertical line corresponding by color to mixing parameters. A mixing parameter of 1 represents LASSO, our baseline model. The variables shown are from the base model shown in Table 4 where the values of the regularization penalty lambda λ are chosen to return 30 predictor variables.

Figure F.8. Correlates of Bureaucrat Effectiveness (Quality): Elastic Net Regularization Coefficients Across Different Mixing Parameters



Note: The figure shows the coefficients from the elastic net regularization procedure on the estimated bureaucrat effects across different values of the mixing parameters. Each coefficient is represented by a small vertical line corresponding by color to mixing parameters. A mixing parameter of 1 represents LASSO, our baseline model. The variables shown are from the base model shown in Table 5 where the values of the regularization penalty lambda λ are chosen to return 30 predictor variables.

Table F.1—: Correlatives of Bureaucrat and Organization Effectiveness: Variable Descriptions

Auctions	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
Auction Length (minutes)	-0.01499 (0.00028)	-0.0056 (0.00101)	0.05	-0.52	3.21	Length of the auction in minutes
Auction Winner Not Chosen	0.00245 (0.00025)	0.00454 (0.00109)	0.02	-0.19	5.33	Indicator if the winner of the auction was ultimately not the suppler listed on the contract
Auction was Held	-0.00081 (0.00026)	0.00092 (0.00079)	0	-10.73	0.09	Indicator if the auction was held (i.e. more than one supplier was admitted to the auction)
Average of Losing Bids / Winning Bid	-0.01006 (0.00028)	-0.00646 (0.001)	0.01	-0.32	26.5	Ratio of the average of all losing bids over the final winning bid
Number of Bidders	-0.01899 (0.00028)	-0.01005 (0.00132)	0.02	-0.77	13.38	Number of bidders that entered bids
Number of Bidders Admitted	-0.03138 (0.00031)	-0.01914 (0.00202)	0.04	-0.73	29.3	Number of bidders admitted to participate in the auctino
Number of Bidders Rejected from Auction	-0.00303 (0.00025)	-0.00417 (0.00125)	0.04	-0.35	50.11	Number of bidders who were not allowed to participate in the auction
Share of Bidders Registered with Tax Authorities	-0.00411 (0.00028)	0.0085 (0.00109)	-0.08	-2.69	0.53	Share of bidders that participated in the auction that were registered with federal tax authorities
Share of Bidders among Firms with High Profit	-0.00228 (0.00028)	0.00675 (0.00141)	-0.11	-1.65	0.86	Share of bidders that participated in the auction that had above-median profits (relative to full sample of suppliers)
Share of Bidders among Firms with High Revenue	-0.00512 (0.00028)	0.00941 (0.0014)	-0.11	-1.83	0.77	Share of bidders that participated in the auction that had above-median revenue (relative to full sample of suppliers)
Share of Bidders from Same County	-0.00942 (0.00028)	0.02388 (0.00154)	-0.02	-0.64	1.94	Share of bidders that participated in the auction that were located in the same county as the End User
Share of Bidders from Same Region	0.01601 (0.00033)	-0.02629 (0.00172)	0.07	-1.53	0.85	Share of bidders that participated in the auction that were located in the same region as the End User
Share of Bidders that are Small Firms	0.00779 (0.00026)	-0.01104 (0.00107)	0.1	-0.41	3.52	Share of bidders that participated in the auction that were registered as small firms
Share of Exporting Bidders	0.00015 (0.00031)	0.01297 (0.00227)	-0.07	-0.2	6.9	Share of bidders that participated in the auction that had exporting activities

Share of Foreign-owned Bidders	0.00012 (0.00028)	-0.00043 (0.00049)	-0.03	-0.16	9.17	Share of bidders that participated in the auction that were foreign-owned
Share of Government Bidders	-0.00447 (0.00036)	0.01322 (0.003)	-0.06	-0.18	7.06	Share of bidders that participated in the auction owned by federal, regional, or municipal governments
Share of Importing Bidders	-0.00392 (0.00031)	0.0043 (0.00192)	-0.08	-0.48	2.75	Share of bidders that participated in the auction that had importing activities
Share of Wholesaler Bidders	0.00481 (0.00027)	0.0075 (0.00166)	0.03	-0.47	2.96	Share of bidders that participated in the auction that operated primarily as wholesale traders
Time between Request and Auction	-0.01778 (3e-04)	0.01337 (0.00295)	-0.04	-2.48	2.19	Number of days elapsed between the day the request was posted and the day the auction was held
Bureaucrats	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
1[Bur. Bought Product Often (Volume)]	-0.00149 (0.00033)		-0.11	-0.51	1.97	Indicator if the main product was also the most common product purchased overall by the Bureaucrat (volume)
1[Bureaucrat Bought from Supplier in Same Year]	-0.00818 (0.00027)		-0.1	-1.22	0.82	Indicator if Supplier won an auction in the previous calendar year with the same bureaucrat
Bureaucrat Product HHI Index (Auctions)	-0.00171 (0.00032)		-0.11	-1.34	3.56	HHI measuring the distribution of auctions (count) by each bureaucrat across two-digit product types
Bureaucrat Product HHI Index (Volume)	0.00042 (0.00027)		0.02	-1.59	3.33	HHI measuring total sales volume of all auctions by each bureaucrat across two-digit product types
Bureaucrat Success Rate	-0.00968 (0.00029)		-0.02	-8.41	1.72	Percentage of requests administered by the Bureaucrat that led to a successful contract
Bureaucrat Supplier HHI Index (Volume)	0.03795 (0.00027)		0.02	-0.82	5.37	HHI measuring total volume of all auctions won by supplier per bureaucrat across two-digit product types
In-house Bureaucrat	0.01191 (0.00036)		0.02	-0.81	1.23	Indicator if the Bureaucrat worked directly at the End User
No. of Auctions Run by Bur. (Auc. Month)	-0.06153 (0.00038)		-0.11	-0.79	3.02	Number of auctions the Bureaucrat was running simultaneously in the same month as the auction
Participation Rate (Bur.)	-0.04297 (0.00041)		0.07	-0.99	52.59	Fraction of the relevant pool of suppliers that Bureaucrat is able to attract to their auction
Participation Rate (Bur., weighted)	-0.0436 (0.00041)		0.07	-1	53.9	Fraction of relevant pool of suppliers that Bureaucrat is able to attract to their auction, weighted by auction volume

Value of Auctions Run by Bur. (Auc. Month)	-0.00053 (0.00032)		-0.11	-5.33	3.77	Total sales volume of the auctions the Bureaucrat was running simultaneously in the same month as the auction
Value of Auctions Run by Bur. (Cumulative, bil. rubles)	-7e-04 (0.00026)		-0.1	-3.76	1.63	Total sales volume of the auctions the Bureaucrat had run cumulatively to the date of the auction
End Users	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
1[End User Bought Product Often (Volume)]		-0.00941 (0.00223)	-0.11	-0.56	1.77	Indicator if the main product was also the most common product purchased overall by the End User (volume)
1[End User Bought from Supplier in Same Year]		0.02738 (0.00177)	-0.13	-1.23	0.81	Indicator if Supplier won an auction in the previous calendar year with the same End User
1[End User works in Agriculture]		0.00373 (0.00087)	0.01	-0.03	30.68	End User works in the agricultural sector
1[End User works in Culture]		0.03958 (0.01356)	0.03	-0.09	11.4	End User works on cultural affairs
1[End User works in Education]		0.02369 (0.00475)	0.15	-0.44	2.25	End User works in education
1[End User works in Emergency Services]		0.00266 (0.00097)	0.02	-0.07	14.06	End User works in emergency services
1[End User works in Environment]		0.01054 (0.01245)	-0.01	-0.1	10.2	End User works in the environmental sector
1[End User works in Forestry]		-0.00397 (0.00081)	0.01	-0.02	41.5	End User works in the forestry sector
1[End User works in Health]		-0.09 (0.00612)	-0.21	-1.44	0.69	End User works in the health care sector
1[End User works in Housing]		0.01436 (0.00157)	0.02	-0.07	14.77	End User works in the housing sector
1[End User works in Internal Affairs]		0.06724 (0.00287)	0.03	-0.11	9.09	End User works in internal affairs (police, justice, etc.)
1[End User works in Labor]		-0.00446 (0.00041)	0.01	-0.03	33.44	End User works in the labor sector (retraining, unemployment assistance, etc.)
1[End User works in Mining]		-0.00037 (8e-05)	0	-0.01	76.8	End User works in the mining sector

1[End User works in Natural Resources]		0.00177 (0.00041)	0	-0.01	182.04	End User works in the natural resources sector
1[End User works in News]		0.00152 (0.00058)	0	-0.01	186.57	End User works in news and journalism
1[End User works in Other ector]		0.03016 (0.00212)	0.08	-0.28	3.63	End User works in other sector
1[End User works in Social Policy]		-0.01921 (7e-04)	0.05	-0.19	5.18	End User works on social policy (welfare, pensions, etc.)
1[End User works in Sport]		0.00859 (0.00105)	0.02	-0.05	19.03	End User works in the sport and recreational sector
1[End User works in Television]		0.00426 (0.00059)	0.01	-0.02	54.13	End User works in television and mass communications
1[End User works in Transportation]		0.01817 (0.00113)	0.02	-0.07	13.5	End User works in the transportation sector
1[End User works in Veterninary Affairs]		0.00491 (0.00164)	0	-0.04	26.76	End User works in veterninary affairs
1[End User works in Youth Services]		0.00499 (0.00077)	0.02	-0.05	20.38	End User works in youth services
Autonomous Organization		-0.01115 (0.00175)	0.01	-0.18	5.58	End User is a non-commercial organization created by the government that enjoys more financial autonomy.
Budget Organization		-0.00645 (0.00107)	0.01	-0.11	9.41	Non-commercial organization with less financial autonomy and stricter budget control from government oversight
Distance from Regional Capital		-0.1025 (0.00295)	0.04	-1.03	1.42	Distance between the End User and the capital of the region where it is located (log kilometers)
End User Average Performance Score		-0.05631 (0.00527)	-0.09	-1.2	2.38	Average performance score across categories for the End User from evaluations by the Federal Treasury
End User Product HHI Index (Auctions)		-0.03466 (0.0046)	-0.07	-1.6	4.33	HHI measuring the distribution of auctions (count) by each End User across two-digit product types
End User Product HHI Index (Volume)		-0.05278 (0.00214)	0.1	-1.47	3.41	HHI measuring total sales volume of all auctions by each end user across two-digit product types
End User Success Rate		0.04045	-0.01	-7.79	1.66	Percentage of requests administered for the End User that led to a successful contract

		(0.00379)				
End User Total Performance Score		0.00198 (0.00106)	-0.08	-1.04	1.81	Total performance score for the End User from independent surveys and evaluations by the Federal Treasury
End Users Supplier HHI Index (Volume)		-0.06991 (0.00353)	0.04	-0.85	5.68	HHI measuring total volume of auctions won by supplier per End User across two-digit product types
Federal Organization		0.24973 (0.00495)	0.04	-0.35	2.87	End User receives funds from the federal government and operates on the federal level
Government Agency		0.00125 (0.00045)	0	-0.03	31.44	End User is classified as a separate government agency, operating more independent of government oversight
Municipal Organization		-0.06477 (0.00389)	0.11	-0.51	1.94	End User receives funds from the municipal government and operates on the municipal level
No. of Auctions Run by End User (Auc. Month)		0.07956 (0.00259)	-0.19	-0.89	2.73	Number of auctions the End User was running simultaneously in the same month as the auction
Other Government Body		0.0115 (0.00161)	-0.02	-4.67	0.21	End User has a much less common legal classification, such as a natural monopoly, audit agency, etc.
Participation Rate (End User)		-0.00742 (0.00315)	0.09	-1.03	38.91	Fraction of the relevant pool of suppliers that End User is able to attract to their auction
Participation Rate (End User, weighted)		-0.00692 (0.00313)	0.09	-1.03	36.42	Fraction of relevant pool of suppliers that End User is able to attract to their auction, weighted by auction volume
Perc. of Auction Volume by End User in Last Calendar Week		-0.03338 (0.00275)	-0.01	-0.88	27.02	Percentage of all auctions (by volume) that End User ran in the last calendar week of the year
Regional Organization		-0.09272 (0.00554)	-0.13	-1.47	0.68	End User receives funds from the regional government and operates on the regional level
Value of Auctions Run by End User (Auc. Month)		0.16843 (0.00371)	-0.14	-6.28	4.17	Total sales volume of the auctions the End User was running simultaneously in the same month as the auction
Value of Auctions Run by End User (Cumulative, bil. rubles)		0.1165 (0.00494)	-0.14	-5.05	2.53	Total sales volume of the auctions the End User had run cumulatively to the date of the auction
Quality	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
1[Contract Term Too Short]	-0.00699 (0.00028)	0.00671 (0.00136)	-0.01	-0.03	30.47	Indicator if amount of time to execute the contract is too short
1[Luxury Product]	-0.00237 (0.00021)	0.00112 (0.00067)	0	-0.01	103.24	Product purchased is considered to be luxury, per data from ClearSpending

Invalid Product Code	0.00254 (0.00021)	-0.00153 (0.00022)	0	-0.01	173.16	Request had an invalid product code per analysis by ClearSpending.Ru
Invalid Product Name	0.01028 (3e-04)	-0.01208 (0.00164)	0.08	-0.44	2.29	Request had an invalid product name per analysis by ClearSpending.Ru
Perc. of Auction Volume by Bur. in Last Calendar Week	-3e-05 (0.00026)		-0.02	-0.73	26.13	Percentage of all auctions (by volume) that Bureaucrat ran in the last calendar week of the year
Regions	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
Public Perceptions of Corruption	-0.85497 (3.14459)	-2.01934 (8.86421)	0	-3.55	1.22	Public perception of the severity of corruption as measured by popular survets
Regional Number of Corruption Cases	-6.56458 (2.30222)	4.32109 (6.02823)	0	-2.71	1.37	Number of corruption cases filed by officials in the region in which the auction was held
Regional Number of Corruption Convictions	-2.40147 (0.65958)	1.67826 (1.44726)	-0.01	-1.59	3.06	Number of corruption convictions secured by officials in the region in which the auction was held
Regional Number of Major Corruption Convictions	-3.36011 (1.02375)	2.20117 (2.39836)	0	-1.63	2.54	Number of major corruption convictions secured by officials in the region in which the auction was held
Regional Number of Officials Found Guilty	-4.9126 (1.156)	4.57753 (1.7449)	0.01	-2.4	1.33	Number of corruption cases where officials were found guilty in the region in which the auction was held
Requests	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
Auction Admit Rate	-0.00013 (0.00026)	0.00308 (0.00136)	-0.02	-5.91	0.4	Percentage of supplier applicants admitted to auction
Deposit (ths. Rubles)	-0.00074 (0.00041)	0.04062 (0.00243)	0.01	-1.07	1.85	Amount bidders are required to deposit before entering auction
No. of Applicants	-0.03036 (0.00031)	-0.0196 (0.00222)	0.05	-0.77	28.49	Number of suppliers that submitted applications to participate in the auction
No. of Products Procured	-0.02496 (0.00036)	-0.00481 (0.00262)	0.09	-0.98	4.1	Number of products overall
Number of 2-Digit Product Codes	-0.00915 (0.00029)	0.00957 (0.00337)	0.17	-0.49	4.61	Number of unique products (as measured by their two-digit codes)
Number of Request Revisions	6e-04 (0.00027)	-0.00773 (0.0013)	0.02	-0.16	21.29	Number of revisions that the Bureaucrat made to the contract before it was finalized

Reservation Price (bil. rubles)	-0.02802 (0.00032)	0.06413 (0.00348)	-0.03	-7.73	5.65	Amount of Reservation price in billions of rubles
Winners	**PwCorr - BurFE**	**PwCorr - OrgFE**	**Mean**	**Min**	**Max**	**Description**
1[Supp. Sold Product Often (Volume)]	-0.01073 (0.00039)	0.00529 (0.0024)	-0.05	-0.89	1.12	Indicator if the main product was also the most common product supplied overall by the Supplier (volume)
1[Supplier Above Median Profit]	0.00241 (0.00026)	0.00281 (0.00121)	-0.09	-1.43	0.7	Indicator if the Supplier has above-median profit relative to the other suppliers in the dataset
1[Supplier Above Median Revenue]	-8e-04 (0.00027)	0.00447 (0.0013)	-0.09	-1.59	0.63	Indicator if the Supplier has above-median revenue relative to the other suppliers in the dataset
1[Supplier from Same Region]	0.01313 (0.00031)	-0.02061 (0.00144)	0.06	-1.4	0.71	Indicator if the Supplier is located in the same region as the End User
1[Supplier is Exporter]	0.00073 (0.00031)	0.0133 (0.00209)	-0.06	-0.17	5.77	Indicator if the Supplier has exporting activities
1[Supplier is Federal Government Agency]	-0.00085 (3e-04)	0.01877 (0.00334)	0.01	-0.06	15.72	Indicator if the Supplier is registered as a federal government agency
1[Supplier is Importer]	-0.00404 (3e-04)	0.00464 (0.00183)	-0.08	-0.43	2.35	Indicator if the Supplier has importing activities
1[Supplier is NGO]	-7e-05 (0.00022)	0.00026 (2e-04)	0.01	-0.02	57.07	Indicator if the Supplier is a nongovernmental organization
1[Supplier is New Firm]	-0.00472 (0.00523)	0.0415 (0.01474)	0.07	-0.43	2.33	Indicator if Supplier is a very new firm
1[Supplier is Private Company]	0.00334 (0.00075)	-0.02762 (0.00526)	-0.04	-2.09	0.48	Indiciator if Supplier is a Private Company
1[Supplier is Regional Government Agency]	-0.00387 (0.00038)	-0.00093 (0.00111)	-0.07	-0.15	6.78	Indicator if the Supplier is registered as a regional government agency
1[Supplier is Wholesale Trader]	0.00687 (0.00026)	0.00485 (0.00144)	0.02	-0.4	2.52	Indicator if Supplier is a wholesale trader
1[Supplier is from Same Postal Code]	-0.00714 (0.00027)	0.02443 (0.00158)	0	-0.59	1.7	Indicator if the Supplier is located in the same postal code as the End User
Supplier - Export Products	0.00285 (0.00031)	0.01004 (0.00219)	-0.07	-0.28	4.04	Number of unique products the Supplier exports
Supplier - Import Products	-0.0052 (0.00031)	0.01508 (0.00232)	-0.07	-0.32	3.73	Number of unique products the Supplier imports

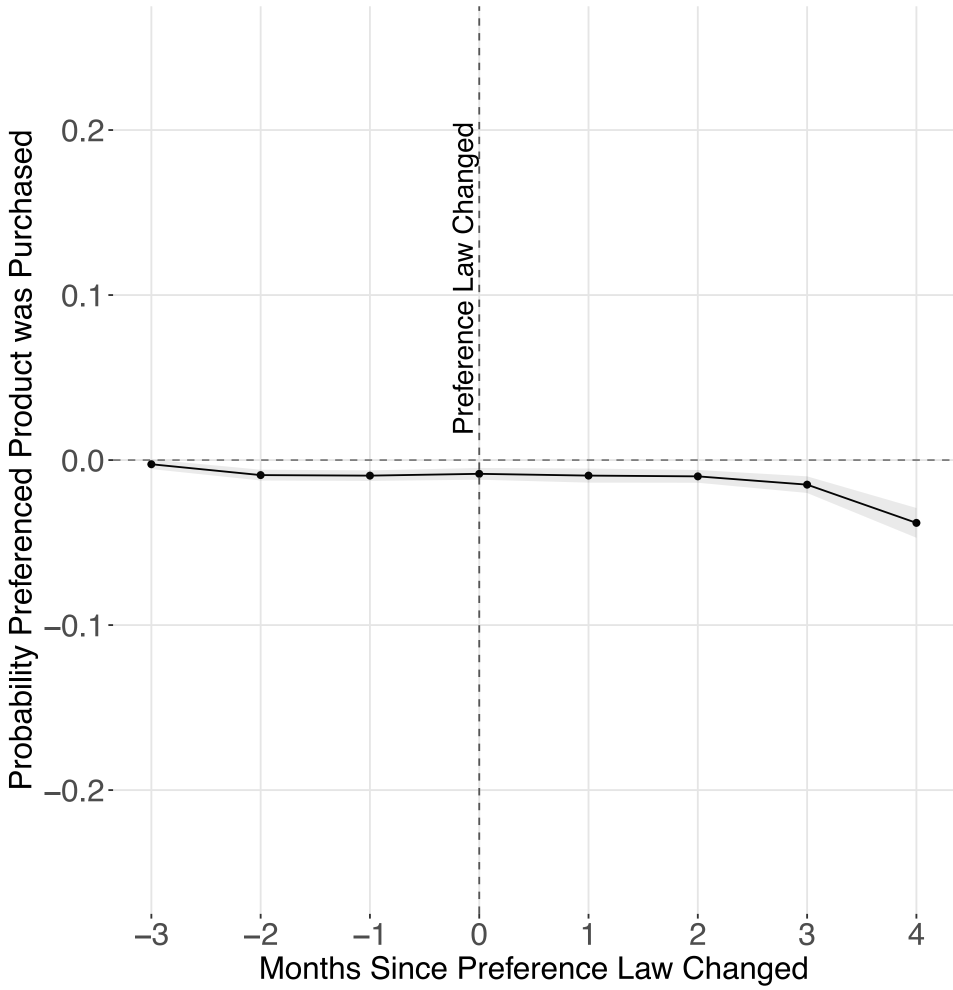
Supplier Age	0.00048 (0.00034)	-7e-04 (0.00154)	-0.08	-1.2	1.88	Age of supplier in years
Supplier All Contracts / Revenue	-0.00287 (0.00028)	0.01297 (0.00176)	0.01	-0.58	2.59	Ratio of Supplier's total contract volume to revenue
Supplier Exports / Revenue	0.00172 (0.00027)	0.00703 (0.00224)	-0.01	-0.06	41.33	Ratio of Supplier's total export volume to revenue
Supplier Imports / Revenue	0.00041 (0.00029)	-0.00415 (0.00114)	-0.02	-0.14	14.12	Ratio of Supplier's total import volume to revenue
Supplier No. of Subsidiaries	-0.00408 (0.00035)	0.00367 (0.00165)	-0.08	-0.3	12.6	Number of subsidiaries owned by Supplier
Supplier Number of Contracts Won	-0.00392 (0.00031)	0.00699 (0.0016)	-0.01	-0.98	2.8	Cumulative number of contracts won by the supplier
Supplier Number of Contracts Won from Large SOEs	-0.00912 (0.00028)	0.00543 (0.0015)	0.01	-0.56	4.41	Cumulative number of contracts won by the supplier under FZ-223 regulating contracts with government agencies
Supplier Number of Countries Exported to	0.00364 (0.00031)	0.01121 (0.00199)	-0.07	-0.29	3.42	Number of unique countries that the Supplier exported to
Supplier Number of Employees	0.00312 (0.00032)	-0.00669 (0.00163)	-0.03	-1.96	4.12	Number of employees working for Supplier
Supplier Product HHI Index (Auctions)	-0.00993 (0.00037)	-0.00791 (0.00401)	-0.12	-2.11	1.61	HHI measuring number of auctions (count) won by supplier across two-digit product types
Supplier Product HHI Index (Volume)	-0.00046 (0.00033)	-0.00497 (0.00287)	-0.03	-2.24	1.55	HHI measuring sales volume of all auctions won by supplier across two-digit product types
Supplier Profit	-0.00148 (0.00033)	0.00726 (0.00188)	-0.11	-4.83	4.16	Supplier net profit
Supplier Profit Per Employee	-0.00811 (3e-04)	0.01117 (0.00199)	-0.08	-0.68	3.14	Ratio of Supplier profits to number of employees
Supplier Revenue (log)	-0.00472 (0.00058)	0.01374 (0.00303)	-0.14	-1.63	3.07	Supplier revenue (log)
Supplier SOE Contracts / Revenue	-0.00225 (0.00026)	0.00873 (0.00177)	0.03	-0.38	3.83	Ratio of Supplier's total volume of contracts with state-owned enterprises to revenue
Supplier Total Assets (log)	-0.0019 (0.00037)	0.01001 (0.00197)	-0.14	-1.11	3.62	Supplier total assets (log)

Supplier Value of Auctions Won (Cumulative, bil. rubles)	-0.00893 (0.00028)	0.00781 (0.00113)	-0.15	-3.94	2.11	Total sales volume of the auctions the Supplier had participated in cumulatively to the date of the auction
Supplier Value of Auctions Won (Total)	-0.00872 (0.00034)	0.0195 (0.00217)	-0.18	-7.8	3.9	Total sales volume of auctions the Supplier was participating in was running simultaneously in the same month
Supplier is Registered with Tax Authorities	0.001 (0.00026)	0.00468 (0.00104)	-0.07	-2	0.5	Indicator if the Supplier is registered with the tax authorities
Supplier is Small Firm	0.00018 (0.00025)	-0.0033 (0.00089)	0.09	-0.33	2.99	Indicator if the Supplier is registered as a small firm
Supplier on Dishonest List	-0.00628 (0.00026)	-0.00315 (0.00087)	0.01	-0.32	3.12	Indicator if Supplier is on the official list of dishonest suppliers

Note: The table describes the full set of variables included in the analysis of bureaucrat and organization effectiveness. The columns 'PwCorr-BurFE' and 'PwCorr-OrgFE' give the pairwise coefficient and standard error between each variable and the estimated bureaucrat and organization effects. Bureaucrat pairwise coefficients are blank for variables not included in the models examining organization effectiveness, while organization pairwise coefficients are blank for variables not included in the models examining bureaucrat effectiveness. Basic summary statistics for each variable are also given, as well as a description of how each was calculated. Firms with less than 100 workers and less than 25 percent ownership by a larger firm do not have to register with the Russian statistical authorities, and are thus not covered by the *Ruslana* data. This includes microenterprises and individual entrepreneurs who participate in procurement and will have missing data. To account for the missing data, we include dummy variables indicating missing data and require the regularization procedure to include them in the final model.

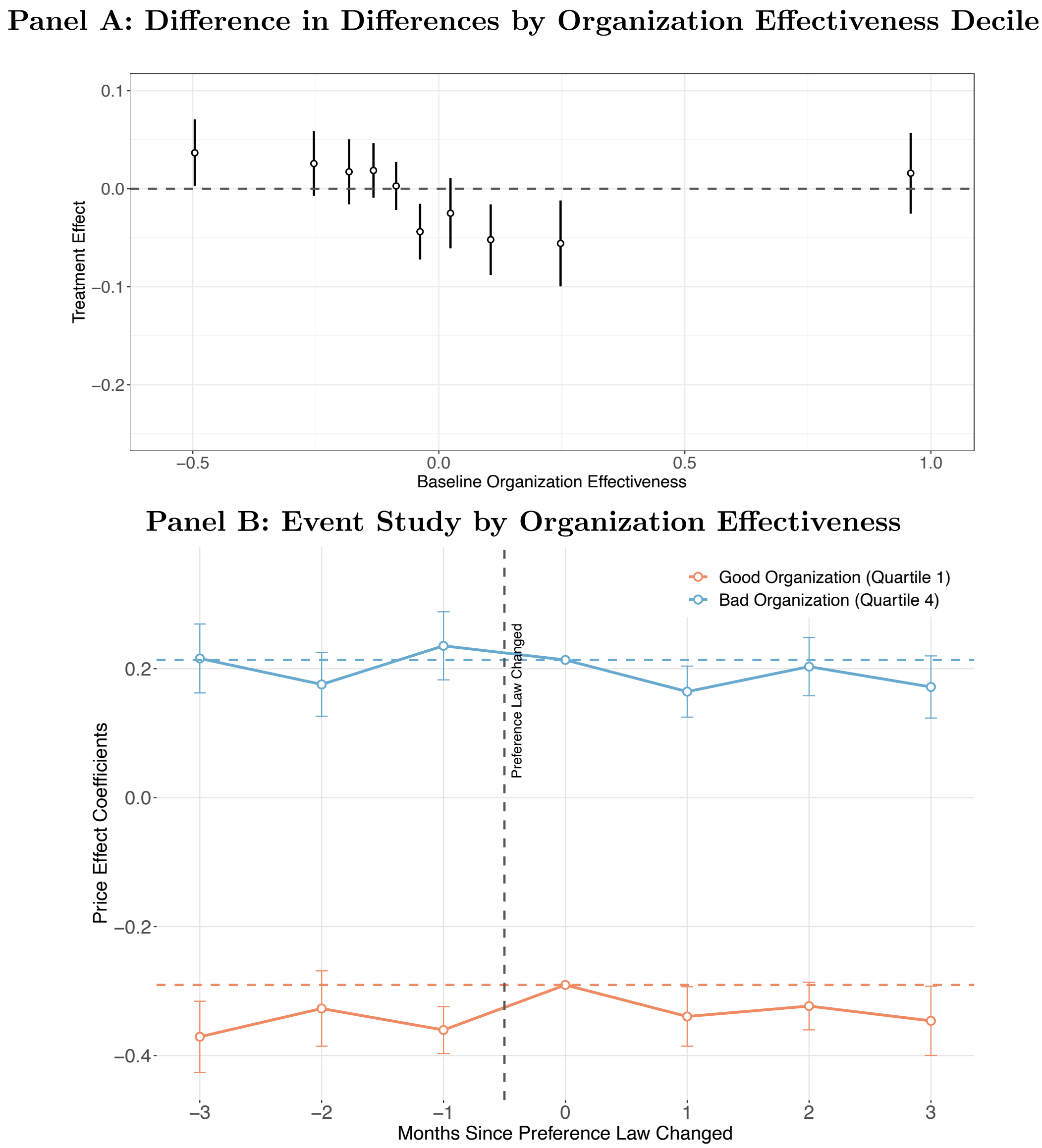
G: Additional Results on Policy Design with a Heterogeneous Bureaucracy

Figure G.1. End Users Do Not Change The Timing of Their Procurement in Anticipation of Preference Laws



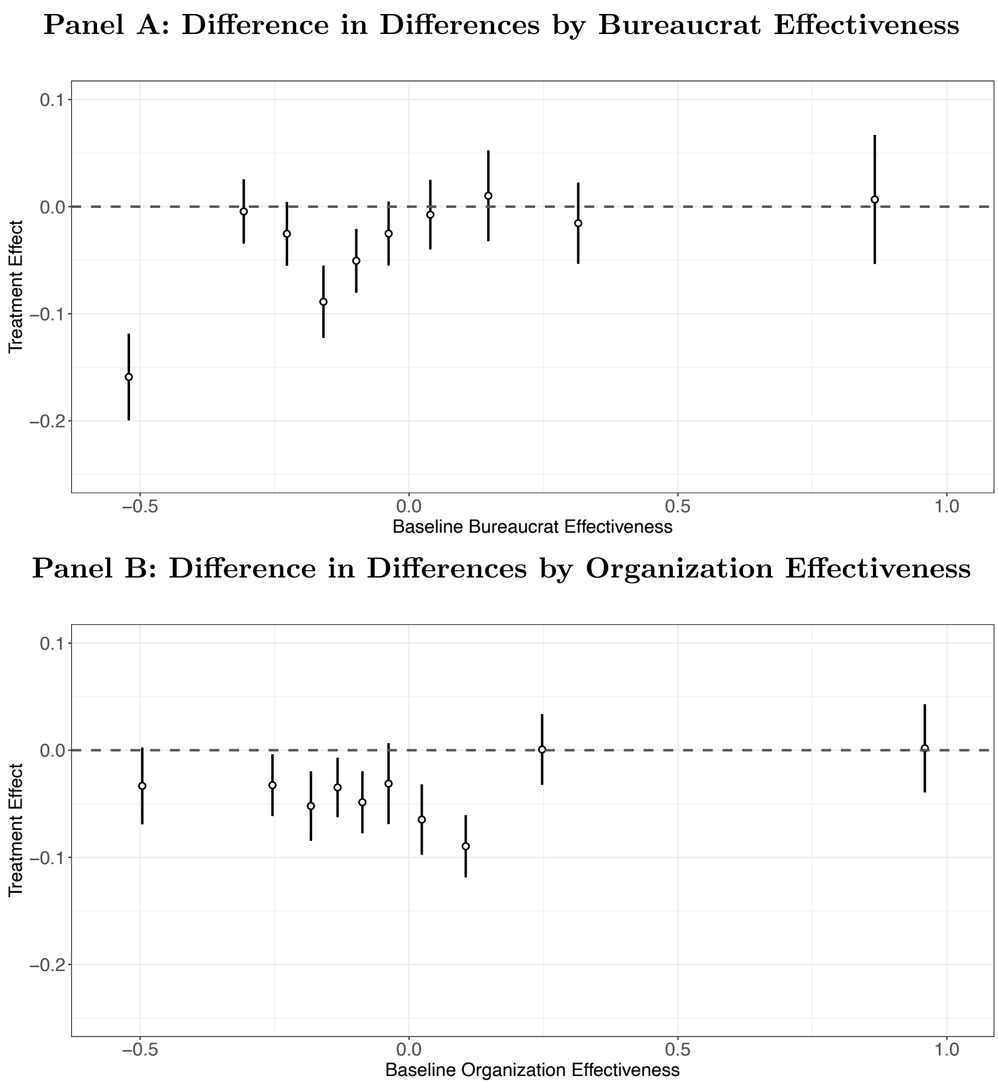
Note: The figure shows the results of an event study analysis of the timing of procurement around the time the preference list is published each year. The x-axis is measured in the number of months preceding or following the activation of the annual preferences laws in 2011, 2012, 2013, and 2014. The dotted vertical lines indicates when the policy was became active. The y-axis in each plot shows the month-specific coefficients from estimation of equation: $Preferenced_{gt} = \mathbf{X}_{igt}\boldsymbol{\beta}+\mu_g+\lambda_t+\mathbf{1}\left\{t-\text{ListMonth}_t = s\right\}+\varepsilon_{igt}$, where Preferenced_{gt} is a dummy indicating that g is on the preferences list in the year month t falls within and ListMonth_t is the month closest to month t in which a preference list is published. $\mathbf{X}_{igt}$ are the same controls we use in Section IV, but we remove the month fixed effects. ε_{igt} is an error term we allow to be clustered by month and good.

Figure G.2. Heterogeneity of Bid Preferences' Effect by Organization Effectiveness



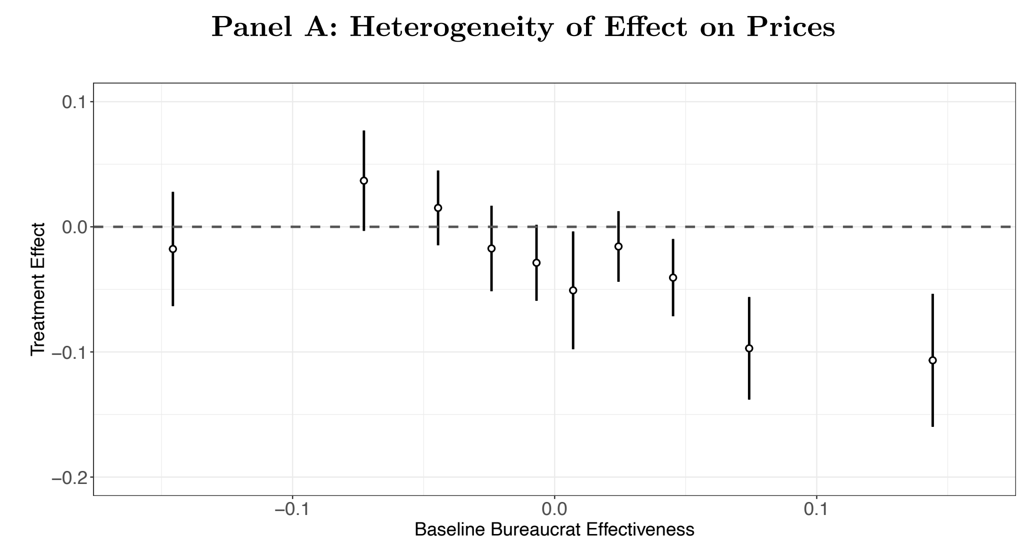
Note: The figure shows how the impacts of the introduction of bid preferences varies by the effectiveness of the implementing organization. Panel A shows estimates from implementing the triple difference model (8) to estimate separate effects for each decile of organization effectiveness: $y_{igt} = \sum_{k=1}^{10} \{D_{kb} + D_{kj} \times (\rho_k \text{Preferenced}_{gt} + \eta_k \text{PolicyActive}_t + \pi_k \text{Preferenced}_{gt} \times \text{PolicyActive}_t)\} + \mathbf{X}_{igt}\beta + \mu_g + \mu_t + \varepsilon_{igt}$ 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 organizations. 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.

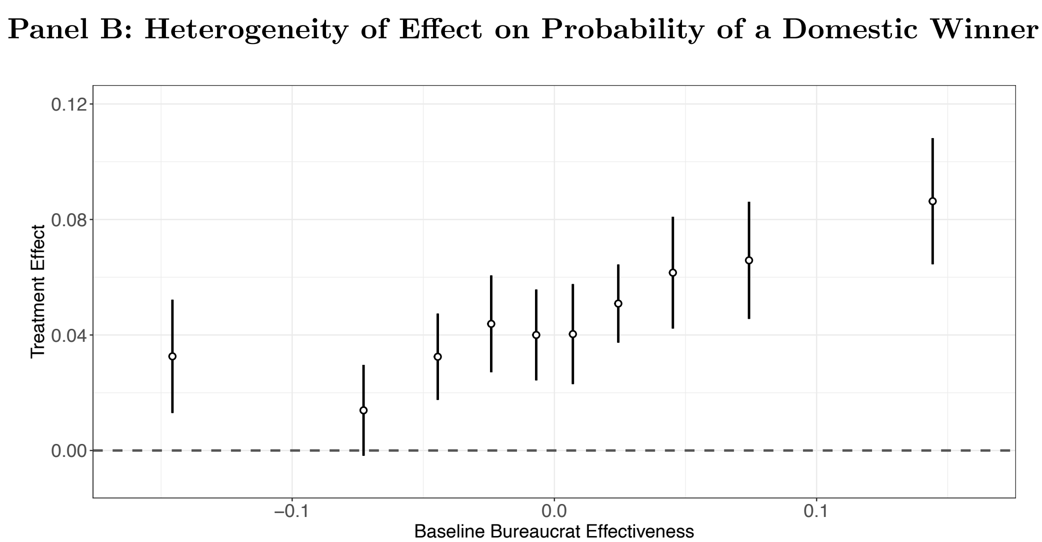
Figure G.3. Heterogeneity of Effect of Bid Preferences on Number of Bidders



Note: The figure shows how the impacts of the introduction of bid preferences on the number of bidders varies by the effectiveness of the implementing buyer. Panel A shows estimates from implementing the triple difference model (8) to estimate separate effects for each decile of bureaucrat effectiveness: $y_{igt} = \sum_{k=1}^{10} \left\{ D_{kj} + D_{kb} \times (\rho_k \text{Preferenced}_{gt} + \eta_k \text{PolicyActive}_t + \pi_k \text{Preferenced}_{gt} \times \text{PolicyActive}_t) \right\} + \mathbf{X}_{igt}\beta + \mu_g + \mu_t + \varepsilon_{igt}$ while Panel B estimates separate effects for each decile of organization effectiveness: $y_{igt} = \sum_{k=1}^{10} \left\{ D_{kb} + D_{kj} \times (\rho_k \text{Preferenced}_{gt} + \eta_k \text{PolicyActive}_t + \pi_k \text{Preferenced}_{gt} \times \text{PolicyActive}_t) \right\} + \mathbf{X}_{igt}\beta + \mu_g + \mu_t + \varepsilon_{igt}$ 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.

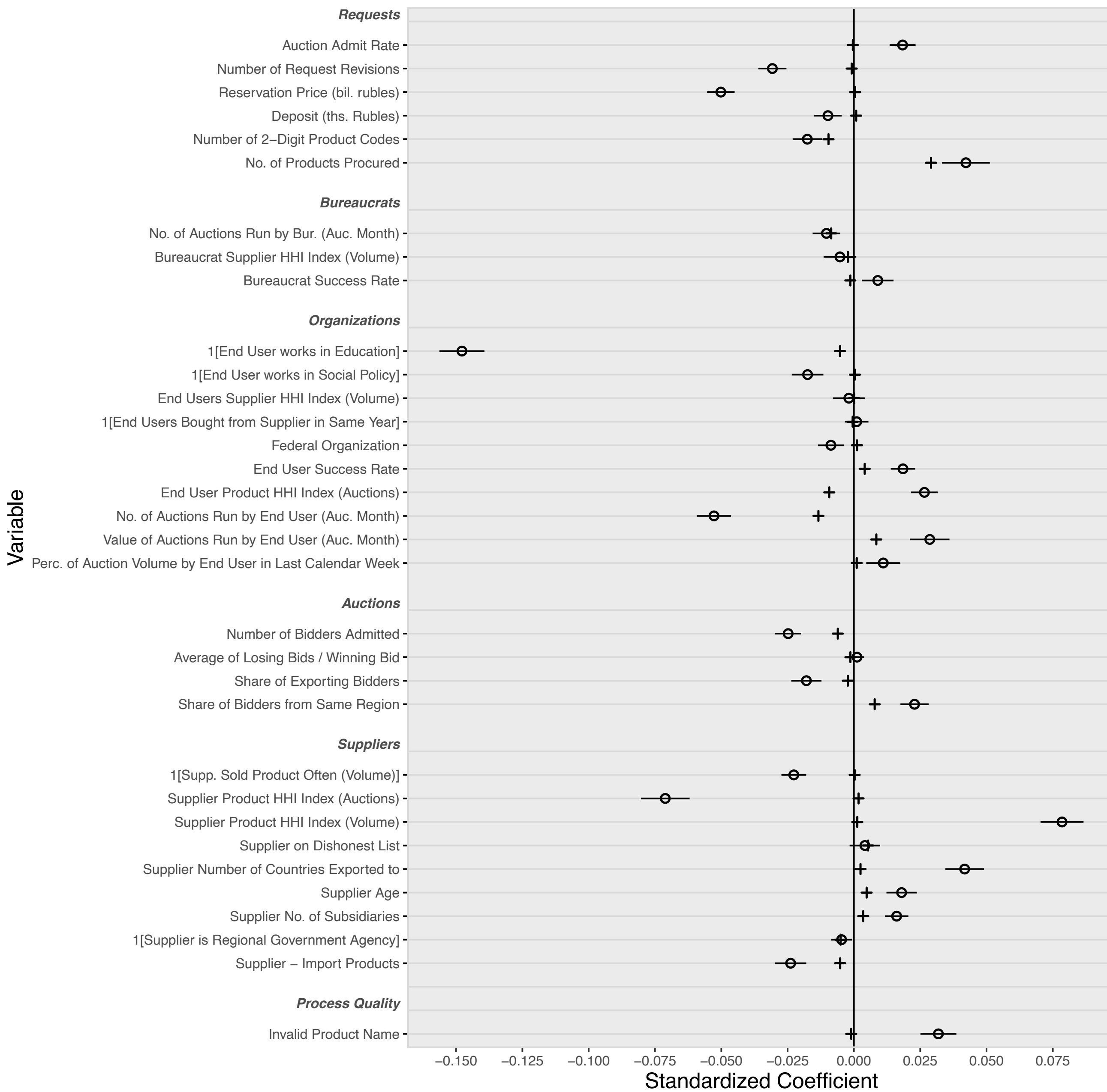
Figure G.4. Heterogeneity of Effect of Bid Preferences in Pharmaceuticals Subsample by Bureaucrat Effectiveness





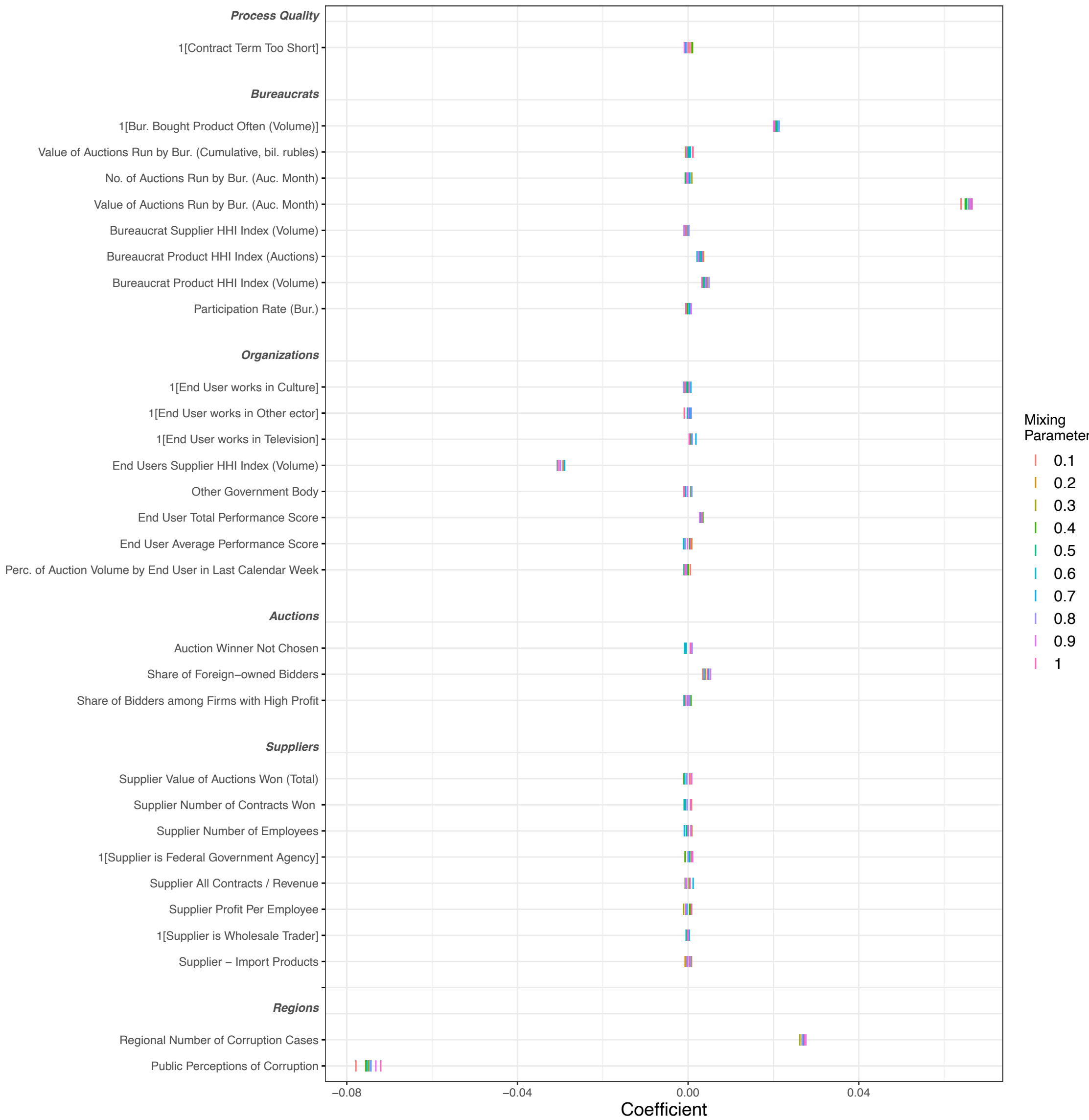
Note: The figure shows how the impacts of the introduction of bid preferences varies by the effectiveness of the implementing buyer in the pharmaceuticals subsample. We estimate the triple difference model (8) to estimate separate effects for each decile of bureaucrat effectiveness: $y_{igt} = \sum_{k=1}^{10} \left\{ D_{kj} + D_{kb} \times (\rho_k \text{Preferenced}_{gt} + \eta_k \text{PolicyActive}_t + \pi_k \text{Preferenced}_{gt} \times \text{PolicyActive}_t) \right\} + \mathbf{X}_{igt}\beta + \mu_g + \mu_t + \varepsilon_{igt}$ 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 A shows effects on prices, while Panel B shows effects on the probability the winning bid offers domestically manufactured pharmaceuticals.

Figure G.5. Predictors of Heterogeneity of Effect of Bid Preferences on Spending Quality



Note: The figure shows the results of estimating our triple-differences specification for heterogeneity of the effect of bid preferences (9): $y_{igt} = \mathbf{X}_{igt}\boldsymbol{\beta} + \mu_g + \lambda_t + \mathbf{Z}_{igt}\boldsymbol{\theta} + \text{Preferenced}_{gt} \times \mathbf{Z}_{igt}\boldsymbol{\gamma} + \text{PolicyActive}_t \times \mathbf{Z}_{igt}\boldsymbol{\eta} + \delta\text{Preferenced}_{gt} \times \text{PolicyActive}_t + \text{Preferenced}_{gt} \times \text{PolicyActive}_t \times \mathbf{Z}_{igt}\boldsymbol{\pi} + \varepsilon_{igt}$ where the elements of the vector of observables $\mathbf{Z}_{igt}$ are picked by LASSO using the largest regularization penalty that returns 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 G.6. Correlates of Price DiD: Elastic Net Regularization Coefficients Across Different Mixing Parameters

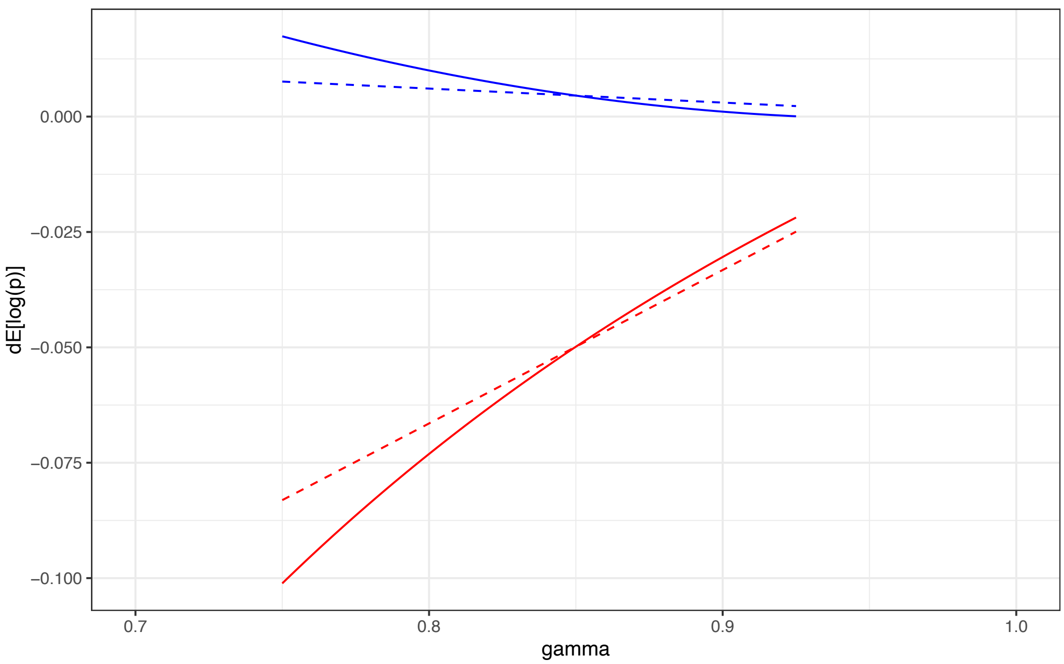


Note: The figure shows the coefficients from the elastic net regularization procedure on the estimated difference-in-differences effects across different values of the mixing parameters. Each coefficient is represented by a small vertical line corresponding by color to mixing parameters. A mixing parameter of 1 represents LASSO, our baseline model. The variables shown are from the base model shown in Figure 8 where the values of the regularization penalty lambda λ is chosen to return 30 variables.

Figure G.7. Correlates of Quality DiD: Elastic Net Regularization Coefficients Across Different Mixing Parameters

Note: The figure shows the coefficients from the elastic net regularization procedure on the estimated difference-in-differences effects across different values of the mixing parameters. Each coefficient is represented by a small vertical line corresponding by color to mixing parameters. A mixing parameter of 1 represents LASSO, our baseline model. The variables shown are from the base model shown in Figure G.5 where the values of the regularization penalty lambda λ are chosen to return 30 variables.

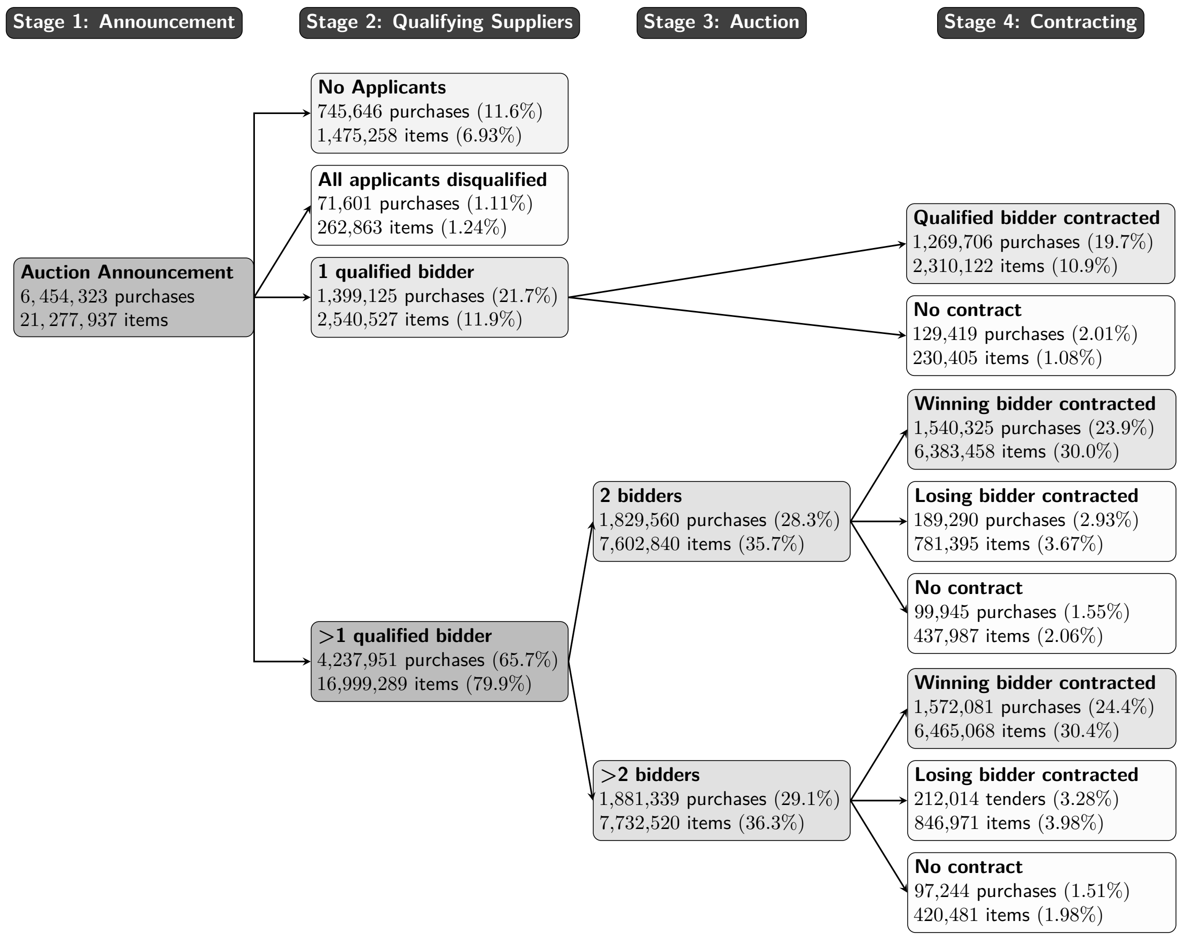
Figure G.8. Constant Semi-Elasticity Approximation for Equivalent Preference Policy Exercise



Note: The figure shows a calibration of the model in section III in solid lines, together with the constant semi-elasticity approximation discussed in section V.D in dashed lines. To calibrate the model we set $\log\left(\bar{\theta}\right) = 1$; the pareto parameter of the productivity distribution of the foreign bidders to be $\delta_F = 1.5$; and the pareto parameter for the local bidders such that the mean productivity is 10% higher for foreign bidders: $\delta_L = 1.588$. We show how the expected log price changes as γ, the fraction of the final bid that a foreign winner receives, changes, as described in proposition 2. The blue lines show this for a high-effectiveness buyer in case 1 of the proposition (specifically, we set $\alpha_c + \psi_c = 0.25$). The red lines show this for a low-effectiveness buyer in case 3 of the proposition (specifically, we set $\alpha_c + \psi_c = 0.85$). The solid and dashed lines are not substantially different from each other.

H: Additional Figures and Tables

Figure H.1. Procurement Process Flow-Chart



Note: This figure lays out the stages of the process public procurement purchases of off-the-shelf goods through electronic auctions follow in Russia. Numbers are based on all purchases made under laws 94 and 44 in 2011-2016. The stages are described in detail in Sub-section I.A.

Table H.1—: Products Covered by Preference Laws, by Year

2011	**2012**	**2013**	**2014**
Live animals	Live animals	Live pigs	Meat and meat products
Textiles	Fresh, chilled, and frozen pork	Fresh, chilled, and frozen pork	Fish and fish products
Clothing and fur products	Sugar	Meat, sausage and other meat products	Salt
Leather and leather goods	Textiles	Cheese, cream and milk	Rice, starches and flour
Chemical products and pharmaceuticals	Clothing and fur products	Rice	Grains, fruits and vegetables (various)
Radio and television equipment	Leather and leather goods	Textiles	Bread, desserts, and chocolate
Medical and measurement equipment	Chemical products and pharmaceuticals	Clothing and fur products	Pharmaceuticals
Cars, trailers and semitrailers	Combine harvesters	Leather and leather goods	Medical and measurement equipment
Transport vehicles (excluding cars)	Self-propelled vehicles	Pharmaceuticals	Ceramic products
	Machinery parts	Agricultural machinery	Iron, steel and ferroalloys (incl. pipes)
	Agricultural machinery	Ratio and television equipment	Steam boilers
	Ratio and television equipment	Medical and measurement equipment	Agricultural machinery
	Medical and measurement equipment	Cars, trailers and semitrailers	Metals and mining equipment
	Cars, trailers and semitrailers	Transport vehicles (excluding cars)	
	Transport vehicles (excluding cars)	Sporting equipment (various)	

Table H.2—Total Procurement in Russia By Type of Mechanism Used

Type	**2011**	%	**2012**	%	**2013**	%	**2014**	%	**2015**	%	**2016**	%	**2011-2016**	%
Electronic Auctions	76.60	46.5	107.65	54.55	106.78	57.98	72.62	51.80	45.13	51.12	45.95	56.39	454.73	53.12
Single Supplier	39.08	23.7	42.95	21.76	39.30	21.34	24.60	17.54	19.61	22.22	19.54	23.98	185.08	21.62
Request for Quotations	6.07	3.7	5.66	2.87	5.32	2.89	1.67	1.19	0.91	1.03	0.77	0.94	20.39	2.38
Open Tender	30.70	18.6	40.86	20.70	32.58	17.69	34.08	24.31	15.82	17.92	10.47	12.85	164.50	19.22
Other Methods	12.17	7.4	0.22	0.11	0.17	0.09	7.23	5.16	6.81	7.72	4.75	5.83	31.36	3.66
Total Procurement	164.62		197.33		184.15		140.19		88.28		81.49		856.06	
Russian Non-Resource GDP	1,720.89		1,873.42		1,989.28		1,786.30		1,231.35		1134.47		9,735.72	
Procurement / Non-Resource GDP (%)	9.6		10.5		9.3		7.8		7.2		7.2		8.8	
Exchange Rate (RUB/USD)	29.37		30.96		31.97		39.20		62.01		66.34		43.31	

Note: This table presents summary statistics about how much procurement was completed under federal laws 94FZ and 44FZ each year according to the mechanism used. All sums are measured in billions of US dollars at current prices using the average ruble-dollar exchange rates shown. Data on Russian procurement comes from the central nationwide Register for public procurement in Russia (`http://zakupki.gov.ru/epz/main/public/home.html`). Data on Russian GDP comes from International Financial Statistics (IFS) at the International Monetary Fund (`http://data.imf.org/`), which we adjust using the percentage of GDP coming from natural resources rents as calculated by the World Bank (`http://data.worldbank.org/indicator/NY.GDP.TOTL.RT.ZS?locations=RU&name_desc=true`).

*

Appendix References

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.

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.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. 2013. *The Elements of Statistical Learning.* Berlin:Springer Series in Statistics.

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

Pang, Sinno Jialin, and Qiang Yang. 2010. "A Survey on Transfer Learning." *IEEE Transactions on Knowledge and Data Engineering*, 22: 1345–1359.

Rifkin, Ryan, and Aldebaro Klautau. 2004. "In Defense of One-Vs-All Classification." *Journal of Machine Learning Research*, 5: 101–141.

Sculley, D. 2010. "Web-Scale K-Means Clustering." *Proceedings of the 19th International Conference on World Wide Web.*

Torrey, Lisa, and Jude Shavlik. 2009. "Transfer Learning." In *Handbook of Research on Machine Learning Applications.* IGI Global.