OA Online Appendix (Not For Publication)

OA.1 Detailed Characterization of Equilibrium Without Bidding Preferences

As shown in section 2.1, the sellers' expected profits can be expressed in terms of their probabilities of winning. Using our assumptions about the distributions of seller fulfillment costs, the probabilities of winning are

$$\begin{aligned}
q_F\left(x;\overline{d}_F,\overline{d}_L\right) &= \mathsf{Pr}\left(b_F\left(x\right)<b_L\left(v_L\right)|v_L\le\overline{d}_L\right)\mathsf{Pr}\left(v_L\le\overline{d}_L\right)+1\times\mathsf{Pr}\left(v_L>\overline{d}_L\right)\\
&= \mathsf{Pr}\left(v_L>x|v_L\le\overline{d}_L\right)\frac{\overline{d}_L-\mu}{1-\mu}+\frac{1-\overline{d}_L}{1-\mu}\\
&= \begin{cases}1 & \text{, if } x<\mu\\ \frac{1-x}{1-\mu} & \text{, if } x\in\left[\mu,\overline{d}_L\right)\\ \frac{1-\overline{d}_L}{1-\mu} & \text{, if } x\ge\overline{d}_L\end{cases} \qquad \text{(OA.1)}\\
q_L\left(x;\overline{d}_F,\overline{d}_L\right) &= \mathsf{Pr}\left(b_L\left(x\right)<b_F\left(v_F\right)|v_F\le\overline{d}_F\right)\mathsf{Pr}\left(v_F\le\overline{d}_F\right)+1\times\mathsf{Pr}\left(v_F>\overline{d}_F\right)\\
&= \mathsf{Pr}\left(v_F>x|v_F\le\overline{d}_F\right)\overline{d}_F+\left(1-\overline{d}_F\right)\\
&= \begin{cases}1-x & \text{, if } x\in\left[\mu,\overline{d}_F\right)\\ 1-\overline{d}_F & \text{, if } x\ge\overline{d}_F\end{cases} \qquad \text{(OA.2)}
\end{aligned}$$

Integrating these probabilities we get the expected profits

$$\begin{aligned}
U_F\left(v;\overline{d}_F,\overline{d}_L\right) &= \int_v^1 q_F\left(x;\overline{d}_F,\overline{d}_L\right)dx\\
&= \begin{cases}\int_v^\mu 1\,dx+\int_\mu^{\overline{d}_L}\frac{1-x}{1-\mu}\,dx+\int_{\overline{d}_L}^1\frac{1-\overline{d}_L}{1-\mu}\,dx & \text{, if } v<\mu\\ \int_v^{\overline{d}_L}\frac{1-x}{1-\mu}\,dx+\int_{\overline{d}_L}^1\frac{1-\overline{d}_L}{1-\mu}\,dx & \text{, if } x\in\left[\mu,\overline{d}_L\right)\\ \int_v^1\frac{1-\overline{d}_L}{1-\mu}\,dx & \text{, if } x\ge\overline{d}_L\end{cases}\\
&= \begin{cases}\frac{2-2\overline{d}_L+\overline{d}_L^2-\mu^2}{2(1-\mu)}-v & \text{, if } v<\mu\\ \frac{2-2\overline{d}_L+\overline{d}_L^2}{2(1-\mu)}-\frac{2v-v^2}{2(1-\mu)} & \text{, if } v\in\left[\mu,\overline{d}_L\right)\\ \frac{(1-\overline{d}_L)(1-v)}{1-\mu} & \text{, if } v\ge\overline{d}_L\end{cases} \qquad \text{(OA.3)}
\end{aligned}$$

And similarly for an entrant of type L with fulfillment cost v (where $\mu < \overline{d}_F$)

$$\begin{aligned} U_L\left(v;\overline{d}_F,\overline{d}_L\right) &= \int_v^1 q_L\left(x;\overline{d}_F,\overline{d}_L\right)dx \\ &= \int_v^{\overline{d}_F}(1-x)\,dx + \int_{\overline{d}_F}^1\left(1-\overline{d}_F\right)dx \\ &= \begin{cases} 1-v-\overline{d}_F+\frac{\overline{d}_F^2}{2}+\frac{1}{2}v^2 & \text{, if } v\in\left[\mu,\overline{d}_F\right) \\ \left(1-\overline{d}_F\right)(1-v) & \text{, if } v\geq\overline{d}_F \end{cases} \end{aligned} \tag{OA.4}$$

To find the entry thresholds, we need to find the type-F supplier $\overline{d}_F$ and type-L supplier $\overline{d}_L$ who are indifferent between entering (in which case they receive $U_i\left(\overline{d}_i;\overline{d}_F,\overline{d}_L\right)-c$) and staying out of the second-stage auction (in which case they receive the contract at price 1 with probability $\frac{1}{2}\left[1-F_j\left(\overline{d}_j\right)\right]$. That is, we need to solve the system of equations

$$\begin{cases} U_F\left(\overline{d}_F;\overline{d}_F,\overline{d}_L\right)-c & =\frac{1}{2}(1-\overline{d}_F)\frac{1-\overline{d}_L}{1-\mu} \\ U_L\left(\overline{d}_L;\overline{d}_F,\overline{d}_L\right)-c & =\frac{1}{2}(1-\overline{d}_F)(1-\overline{d}_L) \end{cases} \tag{OA.5}$$

Since each of these equations has two cases, there are potentially two solutions, depending on whether $\overline{d}_F \lesseqgtr \overline{d}_L$. However, there is no solution when $\overline{d}_F<\overline{d}_L$. The solution with $\overline{d}_F>\overline{d}_L$ satisfies

$$\begin{aligned} &\begin{cases} \frac{(1-\overline{d}_L)(1-\overline{d}_F)}{1-\mu} & =c+\frac{1}{2}(1-\overline{d}_F)\frac{1-\overline{d}_L}{1-\mu} \\ 1-(\overline{d}_L+\overline{d}_F)+\frac{1}{2}\left(\overline{d}_F^2+\overline{d}_L^2\right) & =c+\frac{1}{2}(1-\overline{d}_F)(1-\overline{d}_L) \end{cases} \\ \Leftrightarrow &\begin{cases} 1-\frac{2c(1-\mu)}{1-\overline{d}_L} & =\overline{d}_F \\ \frac{1}{2}(1-\overline{d}_F)(1-\overline{d}_L)+\frac{1}{2\gamma}\left(\overline{d}_F-\overline{d}_L\right)^2 & =c \end{cases} \\ \Leftrightarrow &\begin{cases} 1-\frac{2c(1-\mu)}{1-\overline{d}_L} & =\overline{d}_F \\ \sqrt{2c\mu}+\overline{d}_L & =\overline{d}_F \end{cases} \end{aligned} \tag{OA.6}$$

Solving, we see that

$$\overline{d}_L=\frac{2-\sqrt{2c\mu}-\sqrt{2c\,(2-\mu)}}{2} \tag{OA.7}$$

$$\overline{d}_F=\frac{2+\sqrt{2c\mu}-\sqrt{2c\,(2-\mu)}}{2} \tag{OA.8}$$

which characterize the entry strategies in this equilibrium. Given these, the expected number of entrants in the auction is

$$\mathsf{E}\left[n\right]=G_F\left(\overline{d}_F\right)+G_L\left(\overline{d}_L\right)=\overline{d}_F+\frac{\overline{d}_L-\mu}{1-\mu} \tag{OA.9}$$

we can also calculate the expected payments to each bidder when their fulfillment cost is v

$$
\begin{aligned}
m_F(v) &= U_F(v) + q_F(v)\,v \\
&= \begin{cases} \frac{2-2\overline{d}_L+\overline{d}_L^2-\mu^2}{2(1-\mu)} & \text{, if } v<\mu \\ \frac{2-2\overline{d}_L+\overline{d}_L^2}{2(1-\mu)} - \frac{v^2}{2(1-\mu)} & \text{, if } v\in\left[\mu,\overline{d}_L\right) \\ \frac{(1-\overline{d}_L)}{1-\mu} & \text{, if } v\geq\overline{d}_L \end{cases} \qquad \text{(OA.10)}\\
m_L(v) &= U_L(v) + q_L(v)\,v \\
&= 1-\overline{d}_F+\frac{\overline{d}_F^2}{2}-\frac{1}{2}v^2,\ v\leq\overline{d}_L<\overline{d}_F \qquad \text{(OA.11)}
\end{aligned}
$$

The ex-ante expected profits of the two bidders are therefore

$$
\begin{aligned}
\mathsf{E}_V\left[m_F(v)\right] &= \int_0^{\mu} \frac{2-2\overline{d}_L+\overline{d}_L^2-\mu^2}{2(1-\mu)}\,dv + \int_{\mu}^{\overline{d}_L} \frac{2-2\overline{d}_L+\overline{d}_L^2}{2(1-\mu)} - \frac{v^2}{2(1-\mu)}\,dv + \int_{\overline{d}_L}^{\overline{d}_F} \frac{(1-\overline{d}_L)}{1-\mu}\,dv \\
&= \frac{\overline{d}_L^3-\mu^3+3\overline{d}_F(1-\overline{d}_L)}{3(1-\mu)} \qquad \text{(OA.12)}\\
\mathsf{E}_V\left[m_L(v)\right] &= \int_{\mu}^{\overline{d}_L}\left(1-\overline{d}_F+\frac{\overline{d}_F^2}{2\gamma}-\frac{1}{2}v^2\right)\frac{1}{1-\mu}\,dv = \left[1-\overline{d}_F+\frac{\overline{d}_F^2}{2}\right]\frac{\overline{d}_L-\mu}{1-\mu}+\frac{(\mu^3-\overline{d}_L^3)}{6(1-\mu)} \qquad \text{(OA.13)}
\end{aligned}
$$

Together, these imply that the price the auctioneer expects to pay is

$$
\begin{aligned}
\mathsf{E}\left[p\right] &= \mathsf{E}_V\left[m_F(v)\right]+\mathsf{E}_V\left[m_L(v)\right]+\mathsf{Pr}\left(n=0\right) \\
&= 1-\left(1-\frac{\overline{d}_F}{2}\right)\overline{d}_F\frac{\overline{d}_L-\mu}{1-\mu}+\frac{\overline{d}_L^3-\mu^3}{6\left(1-\mu\right)} \qquad \text{(OA.14)}
\end{aligned}
$$

OA.2 Proof of Proposition 1

Proof. The proposition can be shown by simple differentiation. Starting with the expected number of entrants, differentiating (5), we see that

$$
\frac{\partial\mathsf{E}\left[n\right]}{\partial c} = \frac{\partial\overline{d}_F}{\partial c}+\frac{1}{1-\mu}\frac{\partial\overline{d}_L}{\partial c} \qquad \text{(OA.15)}
$$

which depends on how the entry thresholds change with c. Differentiating the expressions for the entry thresholds (OA.7) and (OA.8),

$$\begin{aligned}\frac{\partial \bar{d}_L}{\partial c} &= \frac{1}{2}\left[-\frac{\sqrt{2c\mu}}{2c} - \frac{\sqrt{2c(2-\mu)}}{2c}\right] \\ &= -\frac{1}{2c}\left(1-\bar{d}_L\right) < 0 \\ \frac{\partial \bar{d}_F}{\partial c} &= \frac{1}{2}\left[\frac{\sqrt{2c\mu}}{2c} - \frac{\sqrt{2c(2-\mu)}}{2c}\right] \\ &= -\frac{1}{2c}\left(1-\bar{d}_F\right) < 0\end{aligned}$$

Plugging these into (OA.15), we obtain

$$\frac{\partial \mathsf{E}[n]}{\partial c} = -\frac{1}{2c}\left[\left(1-\bar{d}_F\right) + \frac{1-\bar{d}_L}{1-\mu}\right] < 0 \tag{OA.16}$$

showing the first part of the proposition. Following the same steps for the second part, the derivative of the expected price is

$$\mathsf{E}[p] = 1 - \left(1-\frac{\bar{d}_F}{2}\right)\bar{d}_F\frac{\bar{d}_L-\mu}{1-\mu} + \frac{\bar{d}_L^3-\mu^3}{6(1-\mu)}$$

and inserting the expressions for the thresholds' derivatives, we obtain

$$\begin{aligned}\frac{\partial \mathsf{E}[p]}{\partial c} &= \frac{1}{2}\frac{\partial \bar{d}_F}{\partial c}\bar{d}_F\frac{\bar{d}_L-\mu}{1-\mu} - \left(1-\frac{\bar{d}_F}{2}\right)\left[\frac{\bar{d}_L-\mu}{1-\mu}\frac{\partial \bar{d}_F}{\partial c} + \frac{\bar{d}_F}{1-\mu}\frac{\partial \bar{d}_L}{\partial c}\right] + \frac{\bar{d}_L^2}{2(1-\mu)}\frac{\partial \bar{d}_L}{\partial c} \\ &= -\frac{\bar{d}_F}{2c}\frac{\bar{d}_L-\mu}{1-\mu}\left(1-\bar{d}_F\right) + \frac{\bar{d}_L-\mu}{1-\mu}\frac{1-\bar{d}_F}{2c} + \left(1-\frac{\bar{d}_F}{2}\right)\frac{\bar{d}_F}{1-\mu}\frac{1-\bar{d}_L}{2c} - \frac{\bar{d}_L^2}{2(1-\mu)}\frac{1-\bar{d}_L}{2c} \\ &= \frac{\bar{d}_L-\mu}{1-\mu}\frac{\left(1-\bar{d}_F\right)^2}{2c} + \frac{1-\bar{d}_L}{4c(1-\mu)}\left[\left(1-\bar{d}_L\right)^2 - \left(1-\bar{d}_F\right)^2 + 2\bar{d}_L\right] > 0\end{aligned}$$

where the last inequality follows since $\bar{d}_F \geq \bar{d}_L$, completing the proof. □

OA.3 Detailed Characterization of Equilibrium With Bidding Preferences

As shown in section 2.1, the sellers' expected profits can be expressed in terms of their probabilities of winning. Using our assumptions about the distributions of seller fulfillment costs, the probabilities of winning are

$$\begin{aligned}
q_F\left(x;\overline{d}_F,\overline{d}_L\right) &= \mathsf{Pr}\left(b_F\left(x\right)<b_L\left(v_L\right)|v_L\le\overline{d}_L\right)\mathsf{Pr}\left(v_L\le\overline{d}_L\right)+1\times\mathsf{Pr}\left(v_L>\overline{d}_L\right)\\
&= \mathsf{Pr}\left(v_L>\frac{x}{\gamma}|v_L\le\overline{d}_L\right)\frac{\overline{d}_L-\mu}{1-\mu}+\frac{1-\overline{d}_L}{1-\mu}\\
&= \begin{cases}1 & \text{, if } x<\gamma\mu\\ \frac{\gamma-x}{\gamma(1-\mu)} & \text{, if } x\in\left[\gamma\mu,\gamma\overline{d}_L\right)\\ \frac{1-\overline{d}_L}{1-\mu} & \text{, if } x\ge\gamma\overline{d}_L\end{cases} \qquad \text{(OA.17)}\\
q_L\left(x;\overline{d}_F,\overline{d}_L\right) &= \mathsf{Pr}\left(b_L\left(x\right)<b_F\left(v_F\right)|v_F\overline{d}_F\right)\mathsf{Pr}\left(v_F\le\overline{d}_F\right)+1\times\mathsf{Pr}\left(v_F>\overline{d}_F\right)\\
&= \mathsf{Pr}\left(\frac{v_F}{\gamma}>x|v_F\le\overline{d}_F\right)\overline{d}_F+\left(1-\overline{d}_F\right)\\
&= \begin{cases}1-x\gamma & \text{, if } x\in\left[\mu,\frac{\overline{d}_F}{\gamma}\right)\\ 1-\overline{d}_F & \text{, if } x\ge\frac{\overline{d}_F}{\gamma}\end{cases} \qquad \text{(OA.18)}
\end{aligned}$$

Integrating these probabilities we get the expected profits

$$\begin{aligned}
U_F\left(v;\overline{d}_F,\overline{d}_L\right) &= \int_v^1 q_F\left(x;\overline{d}_F,\overline{d}_L\right)dx\\
&= \begin{cases}\int_v^{\gamma\mu}1\,dx+\int_{\gamma\mu}^{\gamma\overline{d}_L}\frac{\gamma-x}{\gamma(1-\mu)}\,dx+\int_{\gamma\overline{d}_L}^{1}\frac{1-\overline{d}_L}{1-\mu}\,dx & \text{, if } v<\gamma\mu\\ \int_v^{\gamma\overline{d}_L}\frac{\gamma-x}{\gamma(1-\mu)}\,dx+\int_{\gamma\overline{d}_L}^{1}\frac{1-\overline{d}_L}{1-\mu}\,dx & \text{, if } x\in\left[\gamma\mu,\gamma\overline{d}_L\right)\\ \int_v^1\frac{1-\overline{d}_L}{1-\mu}\,dx & \text{, if } x\ge\gamma\overline{d}_L\end{cases}\\
&= \begin{cases}\frac{2-2\overline{d}_L+\gamma\overline{d}_L^2-\mu^2\gamma}{2(1-\mu)}-v & \text{, if } v<\gamma\mu\\ \frac{2-2\overline{d}_L+\gamma\overline{d}_L^2}{2(1-\mu)}-\frac{2\gamma v-v^2}{2\gamma(1-\mu)} & \text{, if } v\in\left[\gamma\mu,\gamma\overline{d}_L\right)\\ \frac{(1-\overline{d}_L)(1-v)}{1-\mu} & \text{, if } v\ge\gamma\overline{d}_L\end{cases} \qquad \text{(OA.19)}
\end{aligned}$$

And similarly for an entrant of type L with fulfillment cost v (where $\mu<\frac{\overline{d}_F}{\gamma}$)

$$\begin{aligned}
U_L\left(v;\overline{d}_F,\overline{d}_L\right) &= \int_v^1 q_L\left(x;\overline{d}_F,\overline{d}_L\right)dx\\
&= \int_v^{\overline{d}_F/\gamma}\left(1-x\gamma\right)dx+\int_{\overline{d}_F/\gamma}^1\left(1-\overline{d}_F\right)dx\\
&= \begin{cases}1-v-\overline{d}_F+\frac{\overline{d}_F^2}{2\gamma}+\frac{\gamma}{2}v^2 & \text{, if } v\in\left[\mu,\frac{\overline{d}_F}{\gamma}\right)\\ \left(1-\overline{d}_F\right)\left(1-v\right) & \text{, if } v\ge\frac{\overline{d}_F}{\gamma}\end{cases} \qquad \text{(OA.20)}
\end{aligned}$$

To find the entry thresholds, we need to find the type-F supplier $\overline{d}_F$ and type-L supplier $\overline{d}_L$ who are indifferent between entering (in which case they receive $U_i\left(\overline{d}_i;\overline{d}_F,\overline{d}_L\right)-c$) and staying out of the second-stage auction (in which case they receive the contract at price 1 with probability $\frac{1}{2}\left[1-F_j\left(\overline{d}_j\right)\right]$.

That is, we need to solve the system of equations

$$\begin{cases} U_F\left(\overline{d}_F; \overline{d}_F, \overline{d}_L\right) - c & = \frac{1}{2}(1-\overline{d}_F)\frac{1-\overline{d}_L}{1-\mu} \\ U_L\left(\overline{d}_L; \overline{d}_F, \overline{d}_L\right) - c & = \frac{1}{2}(1-\overline{d}_F)(1-\overline{d}_L) \end{cases} \tag{OA.21}$$

Since each of these equations has two cases, there are potentially two solutions, depending on whether $\overline{d}_F \lesseqgtr \gamma \overline{d}_L$. However, there is no solution when $\overline{d}_F < \gamma \overline{d}_L$. The solution with $\overline{d}_F > \gamma \overline{d}_L$ satisfies

$$\begin{aligned} & \begin{cases} \frac{(1-\overline{d}_L)(1-\overline{d}_F)}{1-\mu} & = c + \frac{1}{2}(1-\overline{d}_F)\frac{1-\overline{d}_L}{1-\mu} \\ 1-(\overline{d}_L+\overline{d}_F) + \frac{1}{2\gamma}\left(\overline{d}_F^2 + \gamma^2 \overline{d}_L^2\right) & = c + \frac{1}{2}(1-\overline{d}_F)(1-\overline{d}_L) \end{cases} \\ \Leftrightarrow & \begin{cases} 1 - \frac{2c(1-\mu)}{1-\overline{d}_L} & = \overline{d}_F \\ \frac{1}{2}(1-\overline{d}_F)(1-\overline{d}_L) + \frac{1}{2\gamma}\left(\overline{d}_F - \gamma\overline{d}_L\right)^2 & = c \end{cases} \\ \Leftrightarrow & \begin{cases} 1 - \frac{2c(1-\mu)}{1-\overline{d}_L} & = \overline{d}_F \\ \sqrt{2\gamma c\mu} + \gamma \overline{d}_L & = \overline{d}_F \end{cases} \end{aligned} \tag{OA.22}$$

Solving, we see that

$$\overline{d}_L = \frac{1+\gamma-\sqrt{2\gamma c\mu} - \sqrt{\left[(1-\gamma)-\sqrt{2\gamma c\mu}\right]^2 + 4\gamma c\left(1-\mu\right)}}{2\gamma} \tag{OA.23}$$

$$\overline{d}_F = \frac{1+\gamma+\sqrt{2\gamma c\mu} - \sqrt{\left[(1-\gamma)-\sqrt{2\gamma c\mu}\right]^2 + 4\gamma c\left(1-\mu\right)}}{2} \tag{OA.24}$$

which characterize the entry strategies in this equilibrium. Given these, the expected number of entrants in the auction is

$$\mathsf{E}\left[n\right] = G_F\left(\overline{d}_F\right) + G_L\left(\overline{d}_L\right) = \overline{d}_F + \frac{\overline{d}_L - \mu}{1-\mu} \tag{OA.25}$$

we can also calculate the expected payments to each bidder when their fulfillment cost is v

$$\begin{aligned} m_F\left(v\right) &= U_F\left(v\right) + q_F\left(v\right)v \\ &= \begin{cases} \frac{2-2\overline{d}_L+\gamma\overline{d}_L^2-\mu^2\gamma}{2(1-\mu)} & \text{, if } v < \gamma\mu \\ \frac{2-2\overline{d}_L+\gamma\overline{d}_L^2}{2(1-\mu)} - \frac{v^2}{2\gamma(1-\mu)} & \text{, if } v \in \left[\gamma\mu, \gamma\overline{d}_L\right) \\ \frac{(1-\overline{d}_L)}{1-\mu} & \text{, if } v \geq \gamma\overline{d}_L \end{cases} \end{aligned} \tag{OA.26}$$

$$\begin{aligned} m_L\left(v\right) &= U_L\left(v\right) + q_L\left(v\right)v \\ &= 1 - \overline{d}_F + \frac{\overline{d}_F^2}{2\gamma} - \frac{\gamma}{2}v^2,\ v \leq \overline{d}_L < \overline{d}_F \end{aligned} \tag{OA.27}$$

The ex-ante expected profits of the two bidders are therefore

$$
\begin{aligned}
\mathsf{E}_V\left[m_F(v)\right] &= \int_0^{\gamma\mu} \tfrac{2-2\bar{d}_L+\gamma\bar{d}_L^2-\mu^2\gamma}{2(1-\mu)}\,dv + \int_{\gamma\mu}^{\gamma\bar{d}_L} \tfrac{2-2\bar{d}_L+\gamma\bar{d}_L^2}{2(1-\mu)} - \tfrac{v^2}{2\gamma(1-\mu)}\,dv + \int_{\gamma\bar{d}_L}^{\bar{d}_F} \tfrac{(1-\bar{d}_L)}{1-\mu}\,dv \\
&= \tfrac{\gamma^2(\bar{d}_L^3-\mu^3)+3\bar{d}_F(1-\bar{d}_L)}{3(1-\mu)} \qquad \text{(OA.28)}
\end{aligned}
$$

$$
\mathsf{E}_V\left[m_L(v)\right] = \int_{\mu}^{\bar{d}_L}\left(1-\bar{d}_F+\frac{\bar{d}_F^2}{2\gamma}-\frac{\gamma}{2}v^2\right)\frac{1}{1-\mu}\,dv = \left[1-\bar{d}_F+\frac{\bar{d}_F^2}{2\gamma}\right]\frac{\bar{d}_L-\mu}{1-\mu}+\frac{\gamma(\mu^3-\bar{d}_L^3)}{6(1-\mu)} \qquad \text{(OA.29)}
$$

Together, these imply that the price the auctioneer expects to pay is

$$
\begin{aligned}
\mathsf{E}\left[p\right] &= \mathsf{E}_V\left[m_F(v)\right] + \mathsf{E}_V\left[m_L(v)\right] + \mathsf{Pr}\left(n=0\right) \\
&= 1-\left(1-\frac{\bar{d}_F}{2\gamma}\right)\bar{d}_F\frac{\bar{d}_L-\mu}{1-\mu}+\frac{\bar{d}_L^3-\mu^3}{6\left(1-\mu\right)}\gamma\left(2\gamma-1\right) \qquad \text{(OA.30)}
\end{aligned}
$$

OA.4 Proof of Proposition 2

Proof. We will prove the proposition for the expected number of participants. The proof for the expected price is analogous (but more tedious). To prove the proposition we proceed in three steps. First, we show that for any level of entry costs $c \in (0, \bar{c}]$, there is a threshold γ above which introducing preferences at that rate causes prices to increase, and below which prices decrease. Second, we show that this threshold is decreasing in the entry costs procurers impose on suppliers. Third we argue that these first two steps imply the proposition. Our first step can be characterized in the following lemma.

Lemma 3. *Let $n\left(c,\gamma\right)$ be the expected number of participants when preferences are given by $\gamma \in (0,1]$ and participation costs are $c \in [0,\bar{c}]$. For every $c \in [0,\bar{c}]$ there exists a unique $\gamma^\star\left(c\right) \in [0,1]$ that satisfies $n\left(c,\gamma^\star\left(c\right)\right) = n\left(c,1\right)$. Moreover, $n\left(c,\gamma\right) \leq n\left(c,1\right)\ \forall\gamma \in [0,\gamma^\star\left(c\right)]$ and $n\left(c,\gamma\right) \geq n\left(c,1\right),\ \forall\gamma \in [\gamma^\star\left(c\right),1]$*

Proof. To prove this, we will show that $n\left(c,\gamma\right)$ is unimodal in γ for every c. Differentiating the expected number of entrants, we have that

$$
\frac{\partial n\left(c,\gamma\right)}{\partial\gamma} = \frac{\partial\bar{d}_F}{\partial\gamma} + \frac{1}{1-\mu}\frac{\partial\bar{d}_L}{\partial\gamma}
$$

Denoting the indifference conditions determining the entry thresholds (OA.21) by $\mathcal{F}$, and applying the implicit function theorem, we see that the derivatives are given by

$$
\begin{aligned}
D_\gamma d &= -[D_d\mathcal{F}]^{-1} D_\gamma \mathcal{F} \\
&= -\left|D_d\mathcal{F}\right|^{-1} \begin{pmatrix} \frac{\partial \mathcal{F}_2}{\partial \overline{d}_L} & -\frac{\partial \mathcal{F}_1}{\partial \overline{d}_L} \\ -\frac{\partial \mathcal{F}_2}{\partial \overline{d}_F} & \frac{\partial \mathcal{F}_1}{\partial \overline{d}_F} \end{pmatrix} \begin{pmatrix} \frac{\partial \mathcal{F}_1}{\partial \gamma} \\ \frac{\partial \mathcal{F}_2}{\partial \gamma} \end{pmatrix} \\
&= -\left|D_d\mathcal{F}\right|^{-1} \begin{pmatrix} -\left(1-\gamma\overline{d}_L\right)+\frac{1}{2}\left(1-\overline{d}_F\right) & \frac{1}{2}\frac{1-\overline{d}_F}{1-\mu} \\ 1-\frac{\overline{d}_F}{\gamma}-\frac{1}{2}\left(1-\overline{d}_L\right) & -\frac{1}{2}\frac{1-\overline{d}_L}{1-\mu} \end{pmatrix} \begin{pmatrix} 0 \\ -\frac{\overline{d}_F^2}{2\gamma^2}+\frac{\overline{d}_L^2}{2} \end{pmatrix} \\
&= -\left|D_d\mathcal{F}\right|^{-1} \begin{pmatrix} \frac{1}{2}\frac{1-\overline{d}_F}{1-\mu}\left(\frac{\overline{d}_L^2}{2}-\frac{\overline{d}_F^2}{2\gamma^2}\right) \\ -\frac{1}{2}\frac{1-\overline{d}_L}{1-\mu}\left(\frac{\overline{d}_L^2}{2}-\frac{\overline{d}_F^2}{2\gamma^2}\right) \end{pmatrix}
\end{aligned}
$$

Rearranging the indifference conditions, we can see that $\overline{d}_F = \gamma\overline{d}_L + \sqrt{2\gamma c\mu}$, with which we can show that the determinant in the derivative is

$$
\left|D_d\mathcal{F}\right| = \sqrt{\frac{c\mu}{2\gamma}}\left[\frac{\left(1-\overline{d}_L\right)\gamma+1-\overline{d}_F}{1-\mu}\right] \tag{OA.31}
$$

Substituting in (OA.31) we get that

$$
\begin{pmatrix} \frac{\partial \overline{d}_F}{\partial \gamma} \\ \frac{\partial \overline{d}_L}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} 1-\overline{d}_F \\ -\left(1-\overline{d}_L\right) \end{pmatrix} \frac{\overline{d}_L+\sqrt{\frac{c\mu}{2\gamma}}}{1-\overline{d}_F+\gamma\left(1-\overline{d}_L\right)} \tag{OA.32}
$$

These imply that the derivative of $n(c,\gamma)$ is given by

$$
\frac{\partial n(c,\gamma)}{\partial \gamma} = \left(1-\overline{d}_F-\frac{1-\overline{d}_L}{1-\mu}\right)\left[\frac{\overline{d}_L+\sqrt{\frac{c\mu}{2\gamma}}}{1-\overline{d}_F+\gamma\left(1-\overline{d}_L\right)}\right]
$$

Since the term in square brackets is always positive, the sign of this derivatives depends on the sign of the first term. Since $\partial\overline{d}_F/\partial\gamma > 0$ and $\partial\overline{d}_L/\partial\gamma < 0$, this term is strictly decreasing in γ. Finally, we show that this term is positive at $\gamma = 0$ and negative at $\gamma = 1$. To see that this term is negative at $\gamma = 1$ note that

$$
1-\overline{d}_F-\frac{1-\overline{d}_L}{1-\mu} \leq 1-\overline{d}_F-\left(1-\overline{d}_L\right) = -\left(\overline{d}_F-\overline{d}_L\right) \tag{OA.33}
$$

For any γ, $\overline{d}_F = \gamma\overline{d}_L + \sqrt{2\gamma c\mu}$, so for $\gamma = 1$ we have $\overline{d}_F - \overline{d}_L = \sqrt{2c\mu}$. This implies that

$$
1-\overline{d}_F-\frac{1-\overline{d}_L}{1-\mu} \leq -\sqrt{2c\mu} \leq 0 \tag{OA.34}
$$

where the last inequality is strict whenever c and μ are non-zero. To see that this term is positive at $\gamma = 0$, note that as $\gamma \to 0$, $\overline{d}_F \to 0$. Therefore, to continue to satisfy (OA.21), it must be the case that

$\overline{d}_L \to 1 - 2c\,(1-\mu)$. Therefore,

$$\lim_{\gamma \downarrow 0} \left(1 - \overline{d}_F - \frac{1-\overline{d}_L}{1-\mu}\right) = 1 - 2c > 0 \leftrightarrow c < 1/2 \tag{OA.35}$$

The final ingredient we need to complete the proof of the lemma is to show that there exists exactly one other value of γ for which $\mathsf{E}\,[n]$ is the same as when $\gamma = 1$. To show this, we show that $\mathsf{E}\,[n|\gamma = 0] < \mathsf{E}\,[n|\gamma = 1]$. To see this, note that at $\gamma = 0$ we have that

$$\mathsf{E}\,[n|\gamma = 0] = \overline{d}_F|_{\gamma=0} + \frac{\overline{d}_L|_{\gamma=0} - \mu}{1-\mu} = 1 - 2c \tag{OA.36}$$

At the other limit, when $\gamma = 1$, we get that

$$\begin{aligned}
\overline{d}_F|_{\gamma=1} &= \frac{2 + \sqrt{2c\mu} - \sqrt{2c\mu + 4c\,(1-\mu)}}{2} \\
\overline{d}_L|_{\gamma=1} &= \frac{2 - \sqrt{2c\mu} - \sqrt{2c\mu + 4c\,(1-\mu)}}{2}
\end{aligned}$$

As a result,

$$\begin{aligned}
\mathsf{E}\,[n|\gamma = 1] &= 2 - \frac{\mu}{1-\mu}\frac{\sqrt{2c\mu}}{2} - \frac{2-\mu}{1-\mu}\frac{\sqrt{2c\mu + 4c\,(1-\mu)}}{2} \\
&> 2 - \frac{\mu}{1-\mu}\frac{\sqrt{2c\mu + 4c\,(1-\mu)}}{2} - \frac{2-\mu}{1-\mu}\frac{\sqrt{2c\mu + 4c\,(1-\mu)}}{2} \\
&= 2 - \sqrt{2c\mu + 4c\,(1-\mu)}
\end{aligned}$$

Comparing this to $\mathsf{E}\,[n|\gamma = 0]$, it will be sufficient if

$$\begin{aligned}
2 - \sqrt{2c\mu + 4c\,(1-\mu)} &> 1 - 2c \\
\leftrightarrow 1 + 2c - \sqrt{2c\mu + 4c\,(1-\mu)} &> 0
\end{aligned}$$

Since $\mu \in [0, 1]$,

$$\begin{aligned}
1 + 2c - \sqrt{2c\mu + 4c\,(1-\mu)} &> 1 + 2c - \sqrt{4c} \\
&= 1 + 2\left(c - \sqrt{c}\right) > 1/2 > 0
\end{aligned}$$

Combining all the pieces, $\mathsf{E}\,[n]$ is smaller at $\gamma = 0$ than at $\gamma = 1$ and unimodal in between, so it must have exactly one intermediate $\tilde{\gamma}_n$ for which $\mathsf{E}\,[n|\gamma = \tilde{\gamma}_n] = \mathsf{E}\,[n|\gamma = 1]$, proving the lemma. □

The second step is to show that higher-cost procurers have a lower $\gamma^\star$. The following lemma shows this

Lemma 4. *The price-equalizing γ is lower for procurers who impose larger entry costs on suppliers:*

$$\frac{\partial \gamma^{\star}(c)}{\partial c} < 0 \tag{OA.37}$$

Proof. Applying the implicit function theorem to the expression defining $\gamma^{\star}(c)$, the derivative we are evaluating is given by

$$\frac{\partial \gamma^{\star}}{\partial c} = -\frac{\frac{\partial n(c,\gamma^{\star})}{\partial c} - \frac{\partial n(c,1)}{\partial c}}{\frac{\partial n(c,\gamma^{\star})}{\partial \gamma^{\star}}} \tag{OA.38}$$

By lemma 3, the denominator of (OA.38) is positive, so to show the lemma, we need to show that the numerator is positive. For this, it will be sufficient to show that $\partial^2 n(c,\gamma)/\partial c \partial \gamma$ is negative. To see this, denote the indifference conditions determining the entry thresholds (OA.21) by $\mathcal{F}$ and apply the implicit function theorem. The derivatives of the system with respect to the thresholds $\overline{d}_F$ and $\overline{d}_L$ are in the proof of lemma 3. The remaining derivatives we need are

$$\frac{\partial \mathcal{F}_1}{\partial c} = -1$$
$$\frac{\partial \mathcal{F}_2}{\partial c} = -1$$

Combining all the parts,

$$\begin{aligned}
D_c d &= -[D_d \mathcal{F}]^{-1} D_c \mathcal{F} \\
&= -|D_d \mathcal{F}|^{-1} \begin{pmatrix} -\left(1-\gamma \overline{d}_L\right) + \frac{1}{2}\left(1-\overline{d}_F\right) & \frac{1}{2}\frac{1-\overline{d}_F}{1-\mu} \\ 1 - \frac{\overline{d}_F}{\gamma} - \frac{1}{2}\left(1-\overline{d}_L\right) & -\frac{1}{2}\frac{1-\overline{d}_L}{1-\mu} \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \\
&= -|D_d \mathcal{F}|^{-1} \begin{pmatrix} -\frac{1}{2}\left(1-\overline{d}_F\right) - \sqrt{2\gamma c \mu} & \frac{1}{2}\frac{1-\overline{d}_F}{1-\mu} \\ \frac{1}{2}\left(1-\overline{d}_L\right) - \frac{\sqrt{2\gamma c\mu}}{\gamma} & -\frac{1}{2}\frac{1-\overline{d}_L}{1-\mu} \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \\
&= \sqrt{\frac{2\gamma}{c\mu}} \frac{1-\mu}{\gamma\left(1-\overline{d}_L\right) + \left(1-\overline{d}_F\right)} \begin{pmatrix} \frac{1}{2}\left(1-\frac{1}{1-\mu}\right)\left(1-\overline{d}_F\right) + \sqrt{2\gamma c\mu} \\ -\frac{1}{2}\left(1-\frac{1}{1-\mu}\right)\left(1-\overline{d}_L\right) + \frac{\sqrt{2\gamma c \mu}}{\gamma} \end{pmatrix} \\
&= \begin{pmatrix} \sqrt{\frac{\gamma\mu}{2c}}\left(1-\overline{d}_F\right) - 2\gamma(1-\mu) \\ -\sqrt{\frac{\gamma\mu}{2c}}\left(1-\overline{d}_L\right) - 2(1-\mu) \end{pmatrix} \frac{1}{\gamma\left(1-\overline{d}_L\right) + \left(1-\overline{d}_F\right)}
\end{aligned}$$

Combining these, we see that

$$\begin{aligned}
\frac{\partial \mathsf{E}[n]}{\partial c} &= \frac{\partial \overline{d}_F}{\partial c} + \frac{1}{1-\mu}\frac{\partial \overline{d}_L}{\partial c} \\
&= \frac{1}{1-\overline{d}_F + \gamma\left(1-\overline{d}_L\right)} \left[\sqrt{\frac{\gamma\mu}{2c}}\left(1-\overline{d}_F - \frac{1-\overline{d}_L}{1-\mu}\right) - 2\left[1+\gamma(1-\mu)\right]\right]
\end{aligned}$$

All the terms in the square brackets are decreasing in γ, so we have shown that $\partial^2 n(c,\gamma)/\partial c\partial\gamma$ is negative, and hence we have shown the lemma. □

From these two lemmas the proposition can be seen as follows. To see part (i) consider a particular $\gamma < \overline{\gamma}_n$. By lemma 3, $n(c,\gamma) - n(c,1) > 0$ for all procurers whose entry costs c are such that $\gamma^\star(c) < \gamma$. Conversely, $n(c,\gamma) - n(c,1) < 0$ for all procurers whose entry costs are such that $\gamma^\star(c) > \gamma$. By lemma 4, $\gamma^\star(c) < \gamma$ for all procurers with entry costs higher than $c^\star(\gamma)$, and $\gamma^\star(c) > \gamma$ for all procurers with entry costs below $c^\star(\gamma)$, where $c^\star(\gamma)$ is the unique cost level satisfying $\gamma^\star(c^\star(\gamma)) = \gamma$. Part (ii) follows immediately from the continuity of $n(c,\gamma)$ in c and γ.

□

OA.5 Identification of Bureaucrat and Organization Effects with Multiple Connected Sets

As shown in Abowd *et al.* (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

$$\mathbf{p} = \mathbf{X}\boldsymbol{\beta} + \mathbf{B}\boldsymbol{\alpha} + \mathbf{F}\boldsymbol{\psi} \tag{OA.39}$$

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

$$\begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{F} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\alpha}}_{OLS} \\ \hat{\boldsymbol{\psi}}_{OLS} \end{bmatrix} = \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \end{bmatrix} \mathbf{p} \tag{OA.40}$$

As Abowd *et al.* (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 *et al.* (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

$$\mathbf{p} = \mathbf{X}\boldsymbol{\beta} + \mathbf{B}\tilde{\boldsymbol{\alpha}} + \mathbf{F}\tilde{\boldsymbol{\psi}} + \mathbf{S}\boldsymbol{\gamma} \tag{OA.41}$$

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:

$$\begin{bmatrix} \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \\ \mathbf{S}' \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{F} & \mathbf{S} \end{bmatrix} \\ \begin{bmatrix} \mathbf{S_b} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S_o} & \mathbf{0} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\psi}} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{B}' \\ \mathbf{F}' \\ \mathbf{S}' \end{bmatrix} \mathbf{p} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{OA.42}$$

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 5 (Identification). *If the true model is given by (OA.39), 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 (OA.41) that solve the augmented normal equations (OA.42) (i) are uniquely identified, and (ii) are related to the true bureaucrat and organization effects $\boldsymbol{\alpha}$ and $\boldsymbol{\psi}$ by*

$$\begin{bmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\psi}} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha} - \mathbf{S_b}'\overline{\boldsymbol{\alpha}} \\ \boldsymbol{\psi} - \mathbf{S_o}'\overline{\boldsymbol{\psi}} \\ \overline{\boldsymbol{\alpha}} + \overline{\boldsymbol{\psi}} \end{bmatrix} \tag{OA.43}$$

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 (OA.41) 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 (OA.42) has rank $N_b + N_o + N_s$ and hence the solution to (OA.42) is unique.

To see the second part, it suffices to show that (OA.43) solves (OA.42). First, substitute the estimators out of (OA.42) using (OA.43) and substitute in the true model using (OA.39) to rewrite (OA.42) as

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

From here, noting again that $\mathbf{BS_b}' = \mathbf{FS_o}' = \mathbf{S}$; 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. □

OA.6 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.

OA.6.1 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 section 4.1. Each product description is parsed in the following way, using the Russian libraries for Python's Natural Language Toolkit[59]

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

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]. [60] Similarly, the product description "sodium bicarbonate - solution for infusion 5%,200ml" would result in the following tokens: [sodium, bicarbonate, solution, infusion, 5%, 200ml].[61]

OA.6.2 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 $\mathbf{x}_i \in \mathcal{X}_C$.[62] 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 & Klautau, 2004). 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:

$$\hat{g}_C(\mathbf{x}_i) = \arg\max_{g \in \mathcal{G}_C} p_g(\mathbf{x}_i) \tag{OA.44}$$

Each binary classifier is a linear support vector machine, with a hinge loss function.[63] That is, it solves

$$\min_{\mathbf{w}_g, a_g} \frac{1}{N_C} \sum_{i=1}^{N_C} \max\{0, 1 - y_{gi} \cdot (\mathbf{w}_g \cdot \mathbf{x}_i + a_g)\} \tag{OA.45}$$

[60] 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].

[61] 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].

[62] 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].

[63] A description of the support vector loss function (hinge loss), which estimates the mode of the posterior class probabilities, can be found in Friedman *et al.* (2013, 427)

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 (OA.44). We implement this procedure using the Vowpal Wabbit library for Python.[64] This simple procedure is 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.

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 & Shavlik (2009)). Following the terminology of Pang & 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 OA.1 (translated into English) and OA.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 $\mathcal{X}_C = \mathcal{X}_B$, 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 OA.6.3 below. We use two criteria to identify incorrectly labeled sentences. First, we identify sentences that have been classified as belonging to a certain good, but are very different from the average sentence with that classification in the *customs* data. Second, sentences for which the classifier assigns a low prediction score for all products are deemed to be incorrectly labeled.

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

TABLE OA.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

TABLE OA.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

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 $\mathbf{v}_{gi}$.[65] 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 similarity with the good's mean vector,

$$\theta_{gi} = \frac{\bar{s}_g - s(\mathbf{v}_{gi}, \mathbf{v}_g)}{\bar{s}_g} \tag{OA.46}$$

where $s(\mathbf{v}_{gi}, \mathbf{v}_g) \equiv \cos(\mathbf{v}_{gi}, \mathbf{v}_g) = \frac{\mathbf{v}_{gi}\mathbf{v}_g}{\|\mathbf{v}_{gi}\|\|\mathbf{v}_g\|} = \frac{\sum_{k=1}^{K_g} t_{gik} t_{gk}}{\sqrt{\sum_{k=1}^{K_g} t_{gik}^2}\sqrt{\sum_{k=1}^{K_g} t_{gk}^2}}$ is the cosine similarity of the sentence vector $\mathbf{v}_{gi}$ with its good mean $\mathbf{v}_g$,[66] K_g is the number of tokens used in descriptions of good g, and $\bar{s}_g = \sum_{i=1}^{|\mathbf{t}_g|} s(\mathbf{v}_{gi}, \mathbf{v}_g)$ is the mean of good g's sentence cosine similarities. Sentences with a normalized cosine similarity above a threshold $\overline{\theta}$ are deemed to be misclassified. To choose the threshold $\overline{\theta}$, we use the customs data again. We apply the classification algorithm to the customs data, and identify correctly classified sentences ($\hat{g}_C(\mathbf{x}_i) = g_i$) and incorrectly classified sentences ($\hat{g}_C(\mathbf{x}_i) \neq g_i$). A typical choice of the threshold $\overline{\theta}$ will minimize the sum of type I and type II errors

$$V(\bar{\theta}) = \underbrace{\sum_{\hat{g}_C(\mathbf{x}_i) \neq g_i} I\{\theta_i < \bar{\theta}\}}_{\text{Type I errors}} + \underbrace{\sum_{\hat{g}_C(\mathbf{x}_i) = g_i} I\{\theta_i > \bar{\theta}\}}_{\text{Type II errors}} \tag{OA.47}$$

In the customs data $V(\overline{\theta})$ is roughly flat between 0.65 and 0.95, so we choose 0.95. In our second criterion, we deem a sentence to be incorrectly classified if all predictive scores are below 0.1. i.e. if $\max_{g \in \mathcal{G}_C} p_g(\mathbf{x}_i) < 0.1$.

OA.6.3 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 perform 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

$$\min_{\mathbf{c}} \sum_{i=1}^{N} \|f(\mathbf{c}, \mathbf{t}_i) - \mathbf{t}_i\|^2 \tag{OA.48}$$

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

[65] 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

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

(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

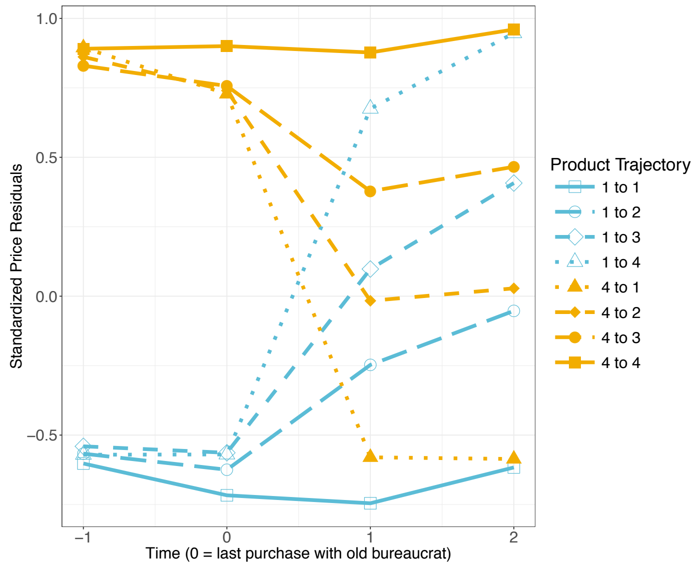
$$\eta\left(i\right)=\frac{b\left(i\right)-a\left(i\right)}{\max\left\{b\left(i\right),a\left(i\right)\right\}} \tag{OA.49}$$

where $a\left(i\right)$ is the average distance between sentence i and the other sentences in the same cluster, and $b\left(i\right)$ 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. Picking $k=10,500$ produces a low silhouette coefficient, and results are not sensitive to using a lower value of $6,500$ or to dropping all the clustered data and using only the correctly classified data.

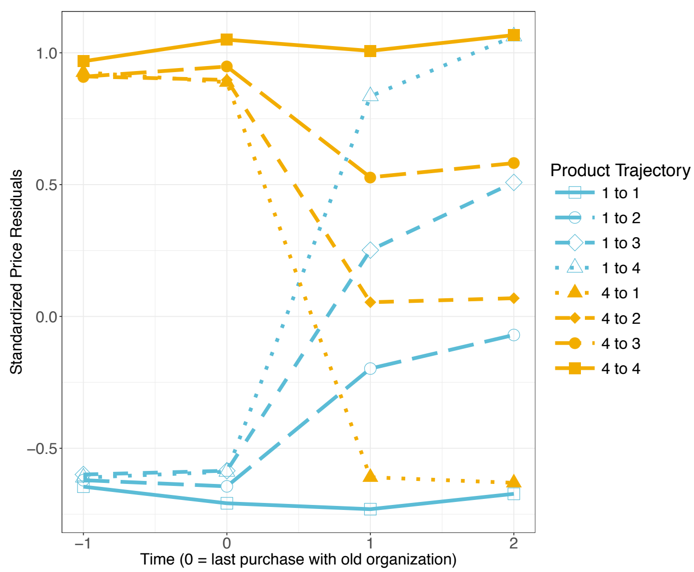
OA.7 Additional Figures and Tables

FIGURE OA.1: EVENT STUDIES OF PRICES AROUND SWITCHES BETWEEN GOODS

PANEL A: EVENT STUDY AROUND BUREAUCRATS SWITCHING GOODS



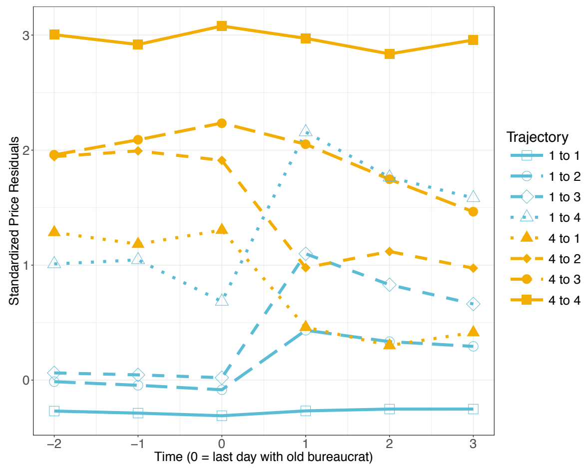
PANEL B: EVENT STUDY AROUND ORGANIZATIONS SWITCHING GOODS



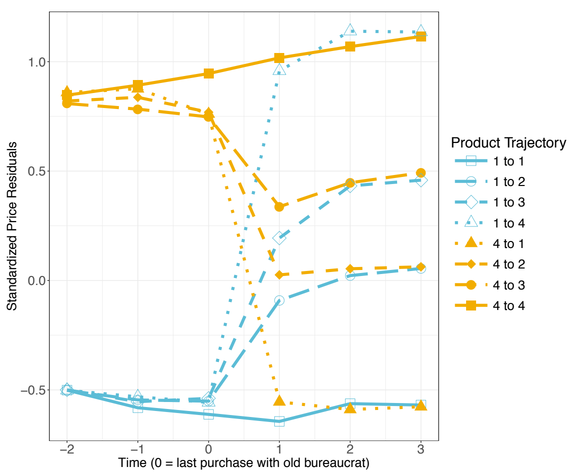
The figure shows time trends in prices around switches in the products that bureaucrats (Panel A) or organizations (Panel B) are purchasing. The horizontal axis measures days on which bureaucrat-product pairs (organization-product pairs in Panel B) occur together, with time 0 being the last day on which the bureaucrat purchases the old product just before switch, and time 1 being the first day the bureaucrat buys the new product after the switch. The y axis measures average residualized prices paid by the bureaucrat-product pair where prices are residualized by regressing log unit prices on month fixed effects. We create a balanced panel in which we require each bureaucrat-product pair to occur together on two separate days and each bureaucrat to purchase at least one other product in the quarter containing time 0 (for the "old" product the bureaucrat purchases before the switch) or time 1 (for the product the "new" product the bureaucrat purchases after the switch). Products are classified into quartiles according to their average (residualized) prices when purchased by other bureaucrats in the quarter containing time 0 (for the old product) or the quarter containing time 1 (for the new product).

FIGURE OA.2: EVENT STUDIES OF PRICES: 3-DAY BALANCED PANELS

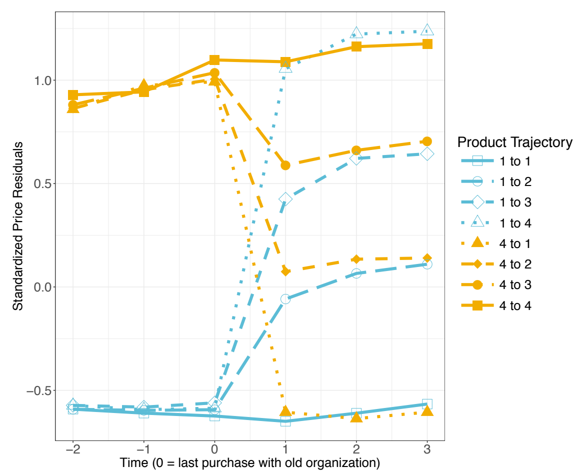
PANEL A: EVENT STUDY AROUND ORGANIZATIONS SWITCHING BUREAUCRATS



PANEL B: EVENT STUDY AROUND BUREAUCRATS SWITCHING GOODS

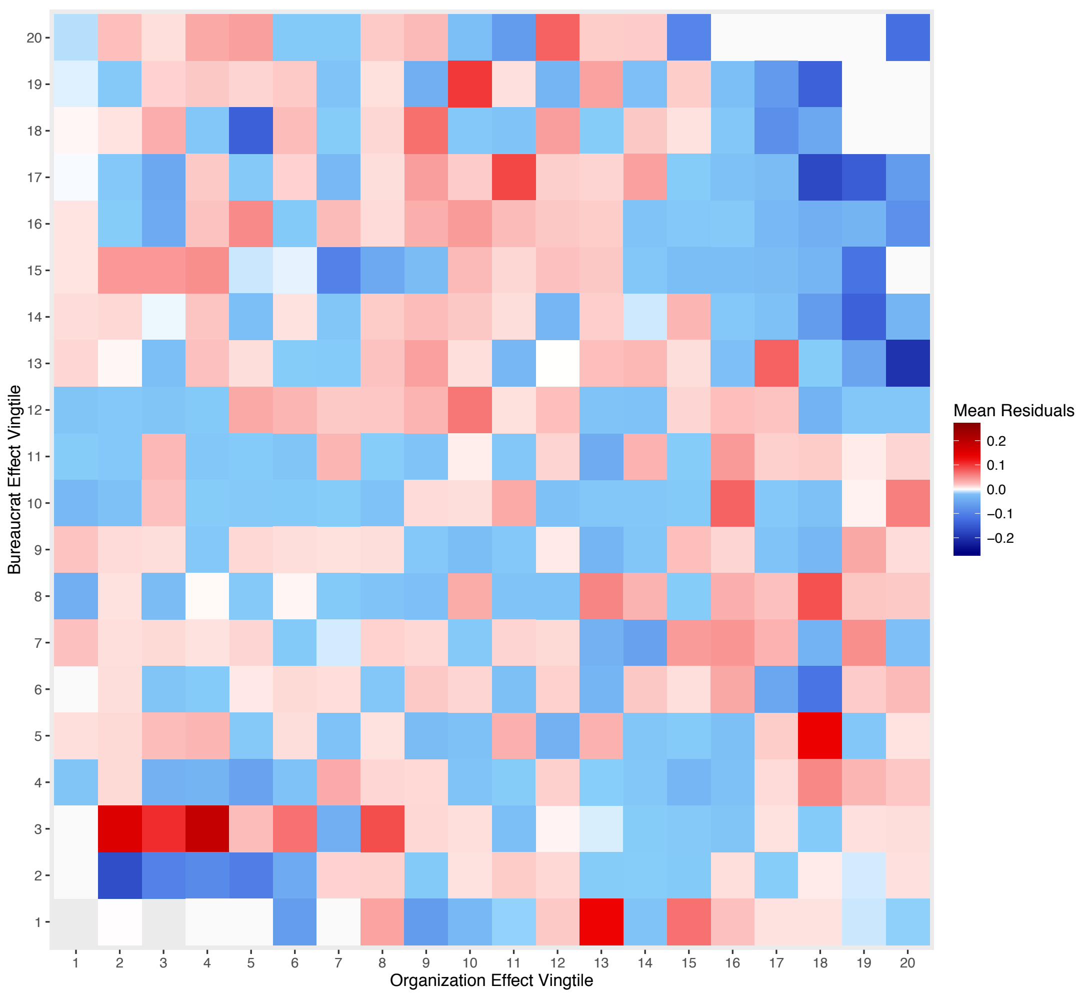


PANEL C: EVENT STUDY AROUND ORGANIZATIONS SWITCHING GOODS



The figure shows time trends in prices around switches in the bureaucrat that organizations use to make purchases (Panel A); the products that bureaucrats are purchasing (Panel B); and the products that organizations are purchasing (Panel C). Panel A is constructed in the same way as figure 2 but with the additional requirement that each bureaucrat-organization pair work together on three separate days. Similarly, Panel B is constructed in the same way as panel A of figure OA.1 but requiring bureaucrat-product pairs to occur on three separate days, and Panel C is constructed in the same way as panel B of figure OA.1 but requiring organization-product pairs to occur on three separate days.

FIGURE OA.3: NO SYSTEMATIC PATTERN IN RESIDUALS: LARGEST CONNECTED SET



The figure presents a heatmap of averages of the residuals from the estimation of equation (11): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. The residuals are binned by vingtiles of the estimated bureaucrat effect $\hat{\alpha}_b$ and organization effect $\hat{\psi}_j$ within each connected set of organizations. The sample used is the Largest Connected Set (All Products) summarized in Table 1.

Table OA.3: Total Procurement in Russia By Type of Mechanism Used

	2011	%	2012	%	2013	%	2014	%	2015	%	2011-2015	%
Electronic Auctions	74.75	46.4	111.12	54.6	113.96	58.0	94.97	51.7	93.66	51.5	488.45	52.7
Single Supplier	38.49	23.9	44.34	21.8	41.87	21.3	32.49	17.7	39.92	21.9	197.12	21.3
Request for Quotations	5.94	3.7	5.81	2.9	5.67	2.9	2.18	1.2	1.88	1.0	21.47	2.3
Open Tender	29.94	18.6	42.10	20.7	34.81	17.7	44.41	24.2	32.64	17.9	183.90	19.8
Other Methods	11.91	7.4	0.20	0.1	0.18	0.1	9.53	5.2	13.85	7.6	35.67	3.8
Total Procurement	161.10		203.64		196.56		183.64		182.02		926.95	
Russian Non-Resource GDP	1,431.68		1,705.01		1,815.10		2,006.63		2,208.35		9,166.77	
Procurement / Non-Resource GDP (%)	11.3		11.9		10.8		9.2		8.2		10.1	

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 an exchange rate of 30 rubles to 1 US dollar. 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).

Table OA.4: Share of Variance of Procurement Prices and Participation explained by Bureaucrats and Organizations: Relaxing Homogeneous Goods Assumption (Khandelwal (2010) Measure)

	(1) Quintile 1	(2) Quintile 2	(3) Quintile 3	(4) Quintile 4	(5) Quintile 5
(1) s.d. of Bur + Org Effects Within CS (across items)	1.005	0.945	0.872	0.809	0.817
(2) s.d. of Total Bur + Org Effects (across items)	1.164	1.119	1.034	0.961	1.008
(3) s.d. of log P	2.621	2.604	2.518	2.378	2.388
(4) s.d. of log P | good, month	1.656	1.683	1.578	1.501	1.526
(5) s.d. of Bur+Org Within Efs / s.d. of log P | good, month	0.607	0.562	0.553	0.539	0.535
(6) s.d. of Bur+Org Total Efs / s.d. of log P | good, month	0.703	0.665	0.655	0.640	0.660
(7) Sample Size	1,411,879	2,831,108	4,271,364	5,727,087	7,055,150

The table implements the variance decomposition in equation (12) using the estimates from equation (11): $p_i = \mathbf{X}_i\boldsymbol{\beta} + \alpha_{b(i,j)} + \psi_j + \gamma_{s(b,j)} + \varepsilon_i$. Each observation is an item procured by an organization j and a bureaucrat indexed by $b(i, j)$. Column (6) uses the sub-sample consisting of all auctions for goods that our text analysis classification method is able to assign a 10-digit product code and that we can match to the scope-for-quality-differentiation ladder developed by Khandelwal (2010). Column (4) removes the quintile with the highest scope-for-quality-differentiation according to the Khandelwal (2010) ladder, Column (3) the highest two quintiles, and so on.

Table OA.5: Average Effect of Bid Preferences for Domestic Producers on Procurement Prices and Auction Entry: Analysis Sample, Raw Fixed Effects

	Prices (P)		Participation (N)	
	Analysis Sample	Largest Connected Set	Analysis Sample	Largest Connected Set
	(1)	(2)	(3)	(4)
log Standardized Quantity	-0.478^{***}	-0.539^{***}	0.026^{***}	0.030^{***}
	(0.023)	(0.017)	(0.003)	(0.003)
Good covered by Prefs.	0.068^{***}	0.043	-0.036	-0.074^{**}
	(0.021)	(0.027)	(0.024)	(0.034)
Policy Active	0.018	0.002	0.229^{**}	0.262^{**}
	(0.025)	(0.042)	(0.107)	(0.122)
Bureaucrat FE	0.945^{***}	0.979^{***}	0.727^{***}	0.830^{***}
	(0.025)	(0.018)	(0.086)	(0.079)
Organization FE	0.952^{***}	0.983^{***}	0.712^{***}	0.813^{***}
	(0.026)	(0.018)	(0.082)	(0.078)
Good covered by Prefs. * Policy Active	-0.115^{***}	-0.134^{***}	-0.117^{***}	-0.168^{***}
	(0.021)	(0.033)	(0.030)	(0.042)
Bureaucrat FE * Good covered by Prefs.	0.036^{**}	0.005	0.061	-0.006
	(0.015)	(0.019)	(0.038)	(0.040)
Bureaucrat FE * Policy Active	-0.005	-0.021	0.010	-0.075
	(0.011)	(0.013)	(0.088)	(0.091)
Organization FE * Good covered by Prefs.	0.045^{***}	0.017	0.077^{**}	0.041
	(0.016)	(0.018)	(0.035)	(0.042)
Organization FE * Policy Active	-0.007	-0.023^{*}	0.032	-0.070
	(0.012)	(0.013)	(0.084)	(0.090)
Bureaucrat FE * Good covered by Prefs. * Policy Active	-0.154^{***}	-0.116^{***}	-0.277^{***}	-0.280^{***}
	(0.020)	(0.025)	(0.051)	(0.058)
Organization FE * Good covered by Prefs. * Policy Active	-0.143^{***}	-0.105^{***}	-0.282^{***}	-0.284^{***}
	(0.020)	(0.023)	(0.051)	(0.064)
Outcome Mean	5.69	6.26	1.64	1.68
Month, Good FEs	Yes	Yes	Yes	Yes
Year×Product×Size×Region FEs	Yes	Yes	Yes	Yes
Connected Set FEs	Yes	Yes	Yes	Yes
Observations	15,957,594	3,973,832	15,957,594	3,973,832
R^2	0.652	0.698	0.377	0.369

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ This table implements a triple-difference approach, interacting the Intent to Treat (ITT) from equation (21) with the estimated bureaucrat and organization effects from Section 5. Unlike 9, the effects included in these models are raw, i.e. they are not estimated using the shrinkage method. In columns (1) and (3) the sample used is the combination of the Analysis Sample summarized in Column (2) of Table 1 and "treated" auctions that the procurers therein carried out. In columns (2) and (4) the sample used is the combination of the Largest Connected Set summarized in Column (3) of Table 1 and "treated" auctions that the procurers therein carried out. The first two columns estimate the triple-difference on the log price paid for each item (P); the second two columns estimate the triple-difference on the number of bidders participating in the auction (N). An item has Preferenced (Good on list) = 1 if the type of good appears on the list of goods covered by the preferences policy for that year. Policy Active = 1 during the part of the relevant year that the preferences policy was in effect. The Outcome Mean is the mean of the dependent variable in the control group, i.e. for goods that were not covered by preferences purchased during the period when the preferences policy was not active. Month and good fixed effects are included in all columns, as are interactions between 2-digit HS Product categories, years, region, and lot size. (We use "product" to distinguish the categories used in these interactions from the much more disaggregate goods categories used for the good fixed effects). Standard errors are clustered on month and good.

Table OA.6: How Effect of Bid Preferences for Domestic Producers on Procurement Prices and Participation varies with Bureaucrat and Organization Effectiveness: Placebo Tests

Panel A: Placebo Test for Moving Policy Active Dates Forward

	Prices (P)				Participation (N)			
	Analysis Sample	1 Month Forward	2 Months Forward	3 Months Forward	Analysis Sample	1 Month Forward	2 Months Forward	3 Months Forward
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bureaucrat FE * Preferenced (Good on list) * Policy Active	−0.183***	−0.151***	−0.122***	−0.121***	−0.279***	−0.233***	−0.221***	−0.172**
	(0.028)	(0.030)	(0.030)	(0.032)	(0.055)	(0.057)	(0.064)	(0.071)
Organization FE * Preferenced (Good on list) * Policy Active	−0.164***	−0.087**	−0.065	−0.081**	−0.307***	−0.282***	−0.247***	−0.246**
	(0.029)	(0.035)	(0.044)	(0.032)	(0.060)	(0.067)	(0.076)	(0.103)
Outcome Mean	5.69	5.44	5.70	5.76	1.64	1.99	1.72	1.81
Month, Good FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Product×Size×Region FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Connected Set FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,957,594	15,957,594	15,957,594	15,957,594	15,957,594	15,957,594	15,957,594	15,957,594
R^2	0.645	0.645	0.645	0.645	0.372	0.370	0.370	0.370

Panel B: Placebo Test for Turning Off Policy in 2015

	Prices (P)				Participation (N)			
	Analysis Sample	3 Months Off	4 Months Off	5 Months Off	Analysis Sample	3 Months Off	4 Months Off	5 Months Off
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bureaucrat FE * Preferenced (Good on list) * Policy Active	−0.183***	−0.017	0.003	0.038	−0.279***	−0.010	−0.011	0.013
	(0.028)	(0.028)	(0.035)	(0.048)	(0.055)	(0.045)	(0.023)	(0.036)
Organization FE * Preferenced (Good on list) * Policy Active	−0.164***	−0.072***	−0.039	−0.021	−0.307***	0.019	0.023	0.009
	(0.029)	(0.024)	(0.035)	(0.044)	(0.060)	(0.061)	(0.035)	(0.037)
Outcome Mean	5.69	5.44	5.70	5.76	1.64	1.99	1.72	1.81
Month, Good FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year×Product×Size×Region FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Connected Set FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,957,594	3,817,427	3,817,427	3,817,427	15,957,594	3,817,427	3,817,427	3,817,427
R^2	0.645	0.650	0.650	0.650	0.372	0.397	0.397	0.397

*** $p<0.01$, ** $p<0.05$, * $p<0.1$ This table implements the same triple-difference approach from Table 9, but includes placebo analysis where the date the preferences policy becomes active is varied. An item has Preferenced (Good on list) = 1 if the type of good appears on the list of goods covered by the preferences policy for that year. In both Panels, columns (1) and (5) are identical to Columns (1) and (2) from Table 9. The main analysis sample and the true date the preference policy became active are used to estimate the triple-difference on the log price paid for each item (P) and the number of bidders participating in the auction (N). In Panel A, the other columns use the main analysis sample but change the date that Policy Active = 1 away from the true date. Columns (2) and (6) move up the dates the preferences became active and went out of effect by 1 month, Columns (3) and (7) by 2 months, etc. In Panel B, the columns (2)-(4) and columns (6)-(8) restrict the sample to only those purchases made in 2015 (when the preferences policy was active throughout). As a placebo test, columns (2) and (5) turn off the preferences policy for the first 3 months of the year, columns (3) and (6) turn off the preferences policy for the first 4 months of the year, etc. Only the estimates of interest are shown (the triple interaction), but all constituent terms and lower interactions are included in the regressions. The Outcome Mean is the mean of the dependent variable in the control group, i.e. for goods that were not covered by preferences purchased during the period when the preferences policy was not active. Month and good fixed effects are included in all columns, as are interactions between 2-digit HS Product categories, years, region, and lot size. (We use "product" to distinguish the categories used in these interactions from the much more disaggregate goods categories used for the good fixed effects). Standard errors are clustered on month and good.