Supplementary Appendix for "Deliver the Vote! Micromotives and Macrobehavior in Electoral Fraud"

This appendix contains proofs of those technical results that do not follow directly from the discussion in the text (A.1-A.5), additional empirical tests for the 2012 Russian presidential election (B.1-B.5), our analysis of the 2011 Russian legislative election (C.1), our analysis of turnout for the 2011 legislative and 2012 presidential Russian elections (D.1), and an analysis of electoral fraud in the 2012 presidential Russian election using Benford's law (E.1).

A.1 The Posterior Density of $\theta | S_i$

Our assumption that θ is uniformly distributed on the unit interval $(0, 1)$ implies that the probability density of θ, $f(\theta)$, is

$$f(\theta) = \begin{cases} 1, & \text{if } 0 < \theta < 1; \\ 0, & \text{otherwise.} \end{cases}$$

Similarly, our assumption that S_i is uniformly distributed on the interval $(\theta - \epsilon, \theta + \epsilon)$ implies that

$$f(S_i|\theta) = \begin{cases} \frac{1}{2\epsilon}, & \text{if } \theta - \epsilon < S_i < \theta + \epsilon; \\ 0, & \text{otherwise.} \end{cases}$$

Using Bayes' rule for random variables, we see that $g(\theta|S_i)$, the posterior density of θ given that agent i observes the incumbent's precinct-level popularity S_i, is

$$g(\theta|S_i) = \frac{f(S_i|\theta)f(\theta)}{f(S_i)} \quad \text{where} \quad f(S_i) = \int_{-\infty}^{\infty} f(S_i|\theta)f(\theta)\, d\theta.$$

Because the support of $f(\theta)$ is limited to $(0, 1)$, while the density $f(S_i|\theta)$ implies that S_i may take values on $(-\epsilon, 1+\epsilon)$, we need to account for the lower and upper bounds 0 and 1 on the integration limits in the computation of $f(S_i)$ when S_i is within an ϵ distance of these bounds. That is, if θ is further than an ϵ distance from the boundaries 0 or 1, $\epsilon < S_i < 1 - \epsilon$, then

$$f(S_i) = \int_{S_i-\epsilon}^{S_i+\epsilon} \frac{1}{2\epsilon}\, d\theta = 1 \quad \text{and} \quad g(\theta|S_i) = \frac{1}{2\epsilon}.$$

Meanwhile, if θ is within an ϵ distance of 0, $-\epsilon < S_i < \epsilon$, then

$$f(S_i) = \int_{0}^{S_i+\epsilon} \frac{1}{2\epsilon}\, d\theta = 1 \quad \text{and} \quad g(\theta|S_i) = \frac{1}{S_i+\epsilon},$$

whereas, if θ is within an ϵ distance of 1, $1-\epsilon < S_i < 1+\epsilon$, then

$$f(S_i) = \int_{S_i-\epsilon}^{1} \frac{1}{2\epsilon}\, d\theta = 1 \quad \text{and} \quad g(\theta|S_i) = \frac{1}{1-(S_i+\epsilon)}.$$

Lemma 1.

$$\theta|S_i \sim \begin{cases} \textit{Uniform}(0, S_i+\epsilon) & \textit{if} \quad -\epsilon < S_i \leq \epsilon; \\ \textit{Uniform}(S_i-\epsilon, S_i+\epsilon) & \textit{if} \quad \epsilon < S_i < 1-\epsilon; \\ \textit{Uniform}(S_i-\epsilon, 1) & \textit{if} \quad 1-\epsilon \leq S_i < 1+\epsilon. \end{cases}$$

Proof. Follows from the text. □

A.2 The Upper Bounds on F and ϵ

Lemma 1 implies that our claim in the main text that the posterior density of $\theta|S_i$ is uniform on the interval $(S_i - \epsilon, S_i + \epsilon)$ holds as long as the signals S_i come from the interval $(\epsilon, 1-\epsilon)$. The interval on which the signals S_i are relevant for the global game analysis in the main text is $(\frac{1}{2} - F - \epsilon, \frac{1}{2} + \epsilon)$. Hence, we must have

$$\epsilon < \frac{1}{2} - F - \epsilon \quad \text{or equivalently} \quad \epsilon < \frac{1}{4}(1 - 2F) \quad \text{and} \quad F < \frac{1}{2}(1 - 4\epsilon).$$

In turn, the maximum admissible values of F and ϵ are

$$\overline{\epsilon} = \frac{1}{4}(1 - 2F) \quad \text{and} \quad \overline{F} = \frac{1}{2}(1 - 4\epsilon),$$

with $\epsilon > 0$ and $F > 0$ implying $\overline{\epsilon} < \frac{1}{4}$ and $\overline{F} < \frac{1}{2}$.

$F < \frac{1}{2}$ is also required for the existence of the left strict-dominance region in our global game. That is, the region $0 < \theta < \frac{1}{2} - F$ in which all agents strictly prefer to refrain from fraud (c.f. **?**, 65).

A.3 Equilibrium reward factor w^*

The incumbent's marginal expected benefit is

$$\frac{\partial[b(1 - \theta^*)]}{\partial w} = \frac{bcF}{(c + w)^2},$$

which is positive and decreasing in w,

$$\frac{\partial^2[b(1 - \theta^*)]}{\partial^2 w} = -\frac{2bcF}{(c + w)^3},$$

Meanwhile, the incumbent's marginal cost is

$$\frac{\partial\left[w\left(\frac{1}{2} + \phi F\right)\right]}{\partial w} = \frac{1}{2(c + w)^2}\left[c^2 + \frac{F^2 + \epsilon + 2F\epsilon}{\epsilon}(2c + 1)w\right],$$

which is also positive but increasing in w,

$$\frac{\partial^2 \left[w\left(\frac{1}{2}+\phi F\right)\right]}{\partial^2 w} = \frac{c^2 F(F+2\epsilon)}{\epsilon(c+w)^3} > 0.$$

Setting the two equal results in a quadratic equation in w,

$$[F+\epsilon(1+2F)]w^2 + 2c(F^2+2\epsilon F+\epsilon)w + \epsilon c(c-2bF) = 0,$$

with the solutions

$$w^* = -c - \sqrt{\frac{cF[cF+2\epsilon(b+c)]}{F^2+2\epsilon F+\epsilon}} \quad \text{and} \quad w^* = -c + \sqrt{\frac{cF[cF+2\epsilon(b+c)]}{F^2+2\epsilon F+\epsilon}}.$$

Of these, only the latter can be positive, which is the case as long as $b > \frac{c}{2F}$.

When $b < \frac{c}{2F}$, the incumbent's marginal cost of fraud is greater than his marginal benefit for any positive value of w. Recall that the incumbent's marginal benefit is positive and decreasing in w with the limiting values

$$\lim_{w\to 0} \frac{\partial[b(1-\theta^*)]}{\partial w} = \frac{bF}{c} \quad \text{and} \quad \lim_{w\to\infty} \frac{\partial[b(1-\theta^*)]}{\partial w} = 0.$$

Meanwhile, the incumbent's marginal cost is also positive but increasing in w with the limiting values

$$\lim_{w\to 0} \frac{\partial \left[w\left(\frac{1}{2}+\phi F\right)\right]}{\partial w} = \frac{1}{2} \quad \text{and} \quad \lim_{w\to\infty} \frac{\partial \left[w\left(\frac{1}{2}+\phi F\right)\right]}{\partial w} = \frac{F^2+2\epsilon F+\epsilon}{2\epsilon}.$$

The latter is larger than the former for any positive value of w if $\lim_{w\to 0} \frac{\partial[b(1-\theta^*)]}{\partial w} < \lim_{w\to 0} \frac{\partial\left[w\left(\frac{1}{2}+\phi F\right)\right]}{\partial w}$, which is the case if $b < \frac{c}{2F}$. For $b \leq \frac{c}{2F}$, therefore, the incumbent's optimal choice of w is $w^* = 0$.

A.4 Comparative Statics

The threshold θ^* is decreasing in w^* since

$$\frac{\partial \theta^*}{\partial w^*} = -\frac{cF}{(c+w^*)^2} < 0.$$

Taking total derivatives of θ^* with respect to the parameters, we see that θ^* is increasing in c and decreasing in b, F, and ϵ:

$$\begin{aligned}
\frac{d\theta^*}{dc} = \frac{\partial \theta^*}{\partial c} + \frac{\partial \theta^*}{\partial w^*}\frac{\partial w^*}{\partial c} &= \frac{w^* F}{(c+w^*)^2} - \frac{cF}{(c+w^*)^2}\left[\frac{F(cF+\epsilon b + 2\epsilon c))}{w^*(F^2+2\epsilon F+\epsilon)} - 1\right] \\
&= \frac{\epsilon b F}{w^*[cF+2\epsilon(b+c)]} > 0; \\
\frac{d\theta^*}{db} = \frac{\partial \theta^*}{\partial b} + \frac{\partial \theta^*}{\partial w^*}\frac{\partial w^*}{\partial b} &= 0 - \frac{cF}{(c+w^*)^2}\frac{\epsilon(c+w^*)}{cF+2\epsilon(b+c)} \\
&= -\frac{\epsilon c F}{(c+w^*)[cF+2\epsilon(b+c)]} < 0; \\
\frac{d\theta^*}{dF} = \frac{\partial \theta^*}{\partial F} + \frac{\partial \theta^*}{\partial w^*}\frac{\partial w^*}{\partial F} &= -\frac{w^*}{c+w^*} - \frac{cF}{(c+w^*)^2}\frac{\epsilon c[\epsilon(b+c) - F(bF-c)]}{(c+w^*)(F^2+2\epsilon F+\epsilon)} \\
&< 0 \text{ for } b > \frac{c}{2F}; \\
\frac{d\theta^*}{d\epsilon} = \frac{\partial \theta^*}{\partial \epsilon} + \frac{\partial \theta^*}{\partial w^*}\frac{\partial w^*}{\partial \epsilon} &= 0 - \frac{cF}{(c+w^*)^2}\frac{cF^2(2bF-c)}{2(c+w^*)(F^2+2\epsilon F+\epsilon)^2} \\
&= -\frac{F(2bF-c)(c+w^*)}{2[cF+2\epsilon(b+c)]^2} < 0 \text{ for } b > \frac{c}{2F}.
\end{aligned}$$

A.5 Fraud as Insurance against Defeat: Pre-election Expectations and the Equilibrium Supply of Fraud

Writing the incumbent's expected payoff from (8) in the main text as

$$\left(\frac{\hat{\theta} + \sigma - \theta^*}{2\sigma}\right) b - w\hat{\theta} - wF\left(\frac{(\hat{\theta} + \epsilon) - S^*}{2\epsilon}\right)$$

and differentiating it with respect to w, we obtain

$$-\frac{b\frac{\partial \theta^*}{\partial w^*}}{2\sigma} - \hat{\theta} - F\left(\frac{(\hat{\theta} + \epsilon) - S^*}{2\epsilon}\right) + \frac{Fw\frac{\partial S^*}{\partial w^*}}{2\epsilon} \tag{A.1}$$

Above, we are treating θ^* and ϕ as functions of w. Their partial derivatives with respect to w are

$$\frac{\partial \theta^*}{\partial w^*} = -\frac{cF}{(c + w^*)^2} \quad \text{and} \quad \frac{\partial \phi}{\partial w^*} = -\frac{cF + 2\epsilon}{(c + w^*)^2}.$$

Substituting these partial derivatives into (A.1) and setting it equal to zero, we obtain a quadratic equation in w. Of the two solutions, only w^* can be positive, which is the case as long as

$$b > \frac{c}{2F}\left(\frac{(2\hat{\theta}F - F + 4\hat{\theta}\epsilon)\sigma}{\epsilon}\right).$$

For $\hat{\theta} = \frac{1}{2}$ and $\sigma = \frac{1}{2}$, w^* reduces to that from the basic model.

A.6 Differences in Competitiveness across Precincts: Persistent versus Variable Electoral Support

The agents' payoffs are summarized in Figure A.1.

		Election result	
		$R \geq \frac{1}{2}$	$R < \frac{1}{2}$
Agent i's action	$a_i = f$	$w(S_i - P_i + F)$	$-cF$
	$a_i = n$	$w(S_i - P_i)$	0

Figure A.1: Agent i's payoffs as a function of her fraud decision a_i and the election result R for the model with persistent and variable electoral support

The majority threshold is now

$$\phi^* = \frac{\frac{1}{2} - \theta - \alpha(\pi - \theta)}{F}.$$

The majority threshold implies that the threshold agent's belief that the incumbent will lose the election is

$$\begin{aligned}
\Pr\left[R < \frac{1}{2} \mid S_i = S^*\right] &= \Pr\left[\phi < \phi^* \mid S_i = S^*\right] \\
&= \Pr\left[\frac{(\theta + \epsilon) - S^*}{2\epsilon} < \phi^*\right] \\
&= \Pr\left[\frac{(\theta + \epsilon) - S^*}{2\epsilon} < \frac{\frac{1}{2} - \theta - \alpha(\pi - \theta)}{F}\right] \\
&= \Pr\left[\theta < \frac{FS^* - \epsilon F + \epsilon - 2\alpha\epsilon\pi}{F + 2\epsilon - 2\alpha\epsilon}\right].
\end{aligned}$$

Given that θ and S_i are uniformly distributed on the intervals $(0, 1)$ and $(\theta - \epsilon, \theta + \epsilon)$, respectively, we have

$$\begin{aligned}
\Pr\left[R < \frac{1}{2} \mid S_i = S^*\right] &= \frac{\frac{FS^* - \epsilon F + \epsilon - 2\alpha\epsilon\pi}{F + 2\epsilon - 2\alpha\epsilon} - (S^* - \epsilon)}{2\epsilon} \\
&= \frac{\frac{1}{2} - (1 - \alpha)(S^* + \epsilon) - \alpha\pi}{F + 2(1 - \alpha)\epsilon}.
\end{aligned}$$

Substituting this expression into the indifference condition in (4), we see that the agents' fraud, popularity, and majority thresholds are

$$
\begin{aligned}
S^* &= \frac{1}{1-\alpha}\left(\frac{1}{2} - F\frac{w}{c+w} - \alpha\pi\right) + \epsilon\frac{c-w}{c+w}\,, \\
\theta^* &= \frac{1}{1-\alpha}\left(\frac{1}{2} - F\frac{w}{c+w} - \alpha\pi\right), \\
\phi^* &= \frac{w}{c+w}\,.
\end{aligned}
$$

Maximizing the incumbent's expected payoff in light of the thresholds S^*, θ^*, and ϕ^* we obtain

$$w^* = \sqrt{\frac{cF(cF + 2\epsilon[b + (1-\alpha)c])}{F^2 + \epsilon(1-\alpha)(1+2F) + \alpha F(\pi - \frac{1}{2})}} - c\,,$$

which is positive as long as

$$b > \frac{c}{2F}\left[\alpha F\left(\pi - \frac{1}{2}\right) + (1-\alpha)\epsilon\right].$$

Otherwise, $w^* = 0$.

Equilibrium uniqueness obtains as long as α and π do not violate the limit dominance condition for global games (c.f. **?**, 65). That is, i) there is $\overline{\theta} \in (0,1)$ such that even if no other agent were to engage in fraud, doing so strictly dominates doing nothing for agent i; ii) there is $\underline{\theta} \in (0,1)$ such that even if all other agents were to engage in fraud, doing nothing strictly dominates engaging in fraud for agent i. Condition i) requires that

$$\alpha\pi + (1-\alpha)\theta \geq \frac{1}{2}, \quad \text{or equivalently that} \quad \theta \geq \frac{\frac{1}{2} - \alpha\pi}{1-\alpha} = \overline{\theta}\,.$$

In order to have $0 < \overline{\theta}$ and $\overline{\theta} < 1$, it must be the case that $\alpha\pi < \frac{1}{2}$ and $\alpha(1-\pi) < \frac{1}{2}$, respectively. Condition ii) requires that

$$\alpha\pi + (1-\alpha)\theta + F < \frac{1}{2}, \quad \text{or equivalently that} \quad \theta < \frac{\frac{1}{2} - F - \alpha\pi}{1-\alpha} = \underline{\theta}\,.$$

In order to have $0 < \underline{\theta}$ and $\underline{\theta} < 1$, it must be the case that $F + \alpha\pi < \frac{1}{2}$ and $\alpha(1-\pi) - F < \frac{1}{2}$, respectively.

Since $F + \alpha\pi < \frac{1}{2}$ implies $\alpha\pi < \frac{1}{2}$ and $\alpha(1-\pi) < \frac{1}{2}$ implies $\alpha(1-\pi) - F < \frac{1}{2}$, equilibrium uniqueness obtains as long as

$$F + \alpha\pi < \frac{1}{2} \quad \text{and} \quad \alpha(1-\pi) < \frac{1}{2}\,. \tag{A.2}$$

These two inequalities hold as long as α is not too large and are most constraining at extreme values of π.[1] Figure A.2 illustrates this requirement by plotting the first inequality in (A.2) by a dashed line, the second inequality in (A.2) by a solid line, and the combinations of α and π that satisfy both inequalities in gray.[2]

A.7 The Normal Model

The Normal information structure outlined in the text implies that the incumbent is genuinely supported by a majority of the electorate, $\theta \geq \frac{1}{2}$, when $\theta' \geq 0$. The fraction of agents ϕ that engage in fraud in equilibrium corresponds to one minus the cumulative distribution function of the $\mathcal{N}(\theta^{*'}, \sigma^2)$ distribution evaluated at $S^{*'}$ and, after observing the incumbent's (probit-transformed) popularity in her precinct S_i', agent i believes that the incumbent's national-level (probit-transformed) popularity θ' follows the Normal density

[1] An alternative reasoning that arrives at these inequalities would look for conditions that rule out equilibria in which agents either always or never commit fraud (i.e. regardless of the value of θ.)

[2] The threshold equilibrium that we examine in the main text still remains one of the multiple equilibria that obtain when the inequalities in (A.2) fail to be satisfied. These inequalities are trivially satisfied when $\alpha = 0$ (our baseline model.)

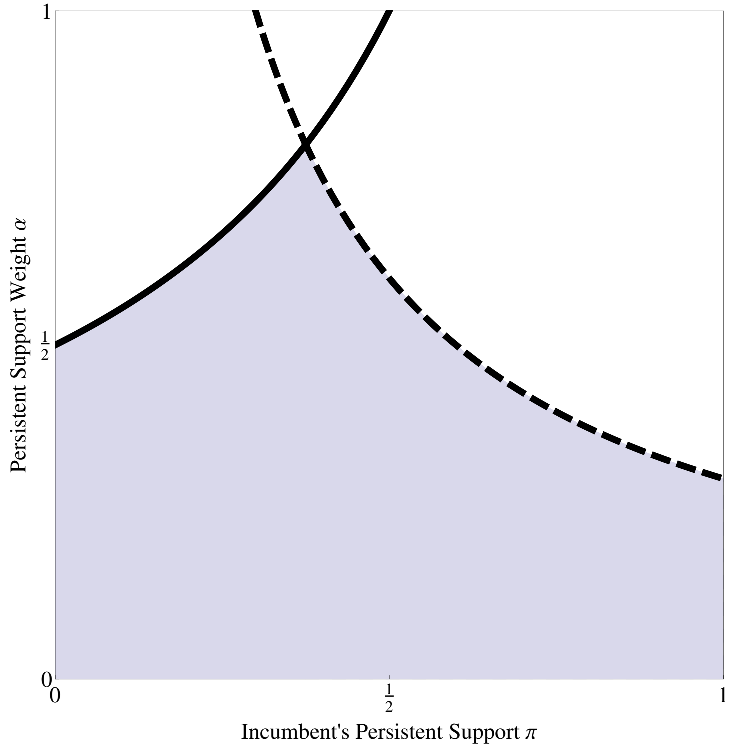


Figure A.2: Values of the incumbent's persistent support π and its weight α that guarantee equilibrium uniqueness

with the mean $\frac{\sigma_0^2 S_i^{'} + \sigma^2 \theta_0^{'}}{\sigma_0^2 + \sigma^2}$ and the variance $\frac{\sigma_0^2 \sigma^2}{\sigma_0^2 + \sigma^2}$.[3]

Figure A.3 illustrates the information structure of the Normal Model by plotting the common prior about the incumbent's popularity θ (dashed line), $\mathcal{N}(\theta_0^{'}, \sigma_0^2)$ with $\theta_0^{'} = 0$, $\sigma_0^2 = 1$, against the threshold agent's posterior belief $\theta | S_i$ about the incumbent's national-level popularity after observing the signal $S_i = S_i^*$ (solid line) for $\sigma^2 = \frac{1}{4}$ (left) and $\sigma^2 = \frac{1}{100}$ (right). The mean $\frac{\sigma_0^2 S_i^{'} + \sigma^2 \theta_0^{'}}{\sigma_0^2 + \sigma^2}$ of the posterior density $\theta | S_i$ is denoted by θ^P.

In order to find the equilibrium fraud, popularity, and majority thresholds, it will be useful to rewrite these quantities as well as the equilibrium conditions in terms of the probit-transformed popularity θ' and the agents' signals $S_i^{'}$. The majority threshold

[3] This is the standard Bayesian inference for the Normal distribution according to which the posterior mean of $\theta' | S_i^{'}$ is a weighted average of the prior mean $\theta_0^{'}$ and the precinct-level signal $S_i^{'}$ (with the weights in proportion of the prior variance σ_0^2 to the signal variance σ^2); see e.g. **?**, 439.

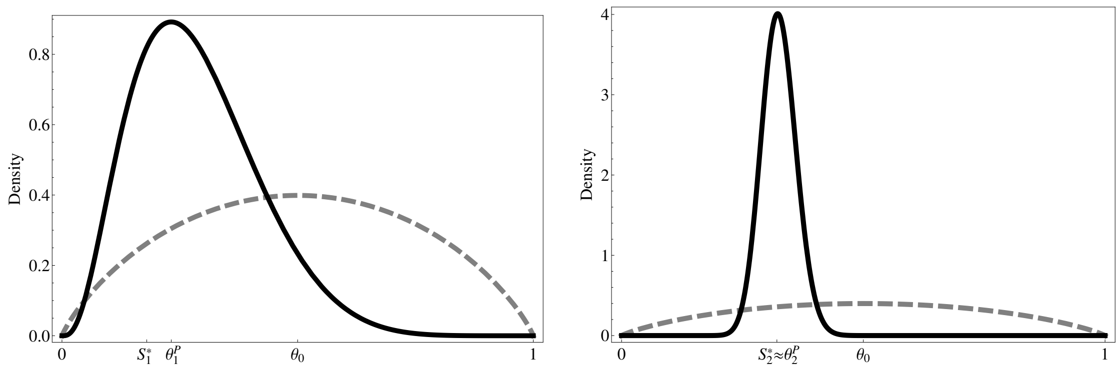


Figure A.3: The common prior $\mathcal{N}(0,1)$ (dashed line) and the threshold agent's posterior belief about the incumbent's national-level popularity after observing the signal $S_i = S_i^*$ (solid line) for $\sigma^2 = \frac{1}{4}$ (left) and $\sigma^2 = \frac{1}{100}$ (right).

becomes

$$\phi^* = \frac{\frac{1}{2} - \Phi(\theta')}{F}; \tag{A.3}$$

the fraction of agents ϕ that engage in fraud in equilibrium corresponds to one minus the cumulative distribution function of the $\mathcal{N}(\theta', \sigma^2)$ distribution evaluated at $S^{*\prime}$,

$$\phi = 1 - \Phi\left(\frac{S^{*\prime} - \theta'}{\sigma}\right); \tag{A.4}$$

and the threshold agent's belief that the incumbent will lose the election (from the indifference condition) becomes

$$\begin{aligned}
\Pr\left[R < \frac{1}{2} \mid S_i = S^*\right] &= \Pr\left[\phi < \phi^* \mid S_i = S^*\right] \\
&= \Pr\left[\phi < \frac{\frac{1}{2} - \Phi(\theta')}{F} \mid S_i = S^*\right] \\
&= \Pr\left[\Phi(\theta') < \frac{1}{2} - \phi F \mid S_i = S^*\right] \\
&= \Pr\left[\theta' < \Phi^{-1}\left(\frac{1}{2} - \phi F\right) \mid S_i = S^*\right],
\end{aligned}$$

which corresponds to the cumulative distribution function of the $\mathcal{N}\left(\frac{\sigma_0^2 S_i^{'}+\sigma^2\theta_0^{'}}{\sigma_0^2+\sigma^2}, \frac{\sigma_0^2\sigma^2}{\sigma_0^2+\sigma^2}\right)$ distribution evaluated at $\Phi^{-1}\left(\frac{1}{2}-\phi F\right)$,

$$\Pr\left[R<\frac{1}{2}\mid S_i=S^*\right]=\Phi\left(\frac{\Phi^{-1}\left(\frac{1}{2}-\phi F\right)-\frac{\sigma_0^2 S_i^{'}+\sigma^2\theta_0^{'}}{\sigma_0^2+\sigma^2}}{\sqrt{\frac{\sigma_0^2\sigma^2}{\sigma_0^2+\sigma^2}}}\right). \tag{A.5}$$

Updating the indifference condition using (A.5) and combining with (A.4), we see that in equilibrium the following two conditions must hold for the Normal model:

$$\phi = \frac{\frac{1}{2}-\Phi(\theta^{*'})}{F} \quad \text{and,} \tag{A.6}$$

$$\Phi\left(\frac{\Phi^{-1}\left(\frac{1}{2}-\phi F\right)-\frac{\sigma_0^2 S^{*'}+\sigma^2\theta_0^{'}}{\sigma_0^2+\sigma^2}}{\sqrt{\frac{\sigma_0^2\sigma^2}{\sigma_0^2+\sigma^2}}}\right) = \frac{w}{c+w}, \tag{A.7}$$

with $\phi = 1-\Phi\left(\frac{S^{*'}-\theta^{*'}}{\sigma}\right)$ according to (A.4). Equilibrium condition (A.6) states that, when $S_i^{'}=S^{*'}$ and $\theta^{'}=\theta^{*'}$, the fraction of agents that receive a signal of at least $S^{*'}$ is exactly the fraction of agents needed to deliver a bare majority of the vote to the incumbent. Equilibrium condition (A.7) states that the threshold agent with the signal $S_i^{'}=S^{*'}$ is indifferent between engaging in fraud and refraining from it.

This set of two equations about two unknowns can be reduced to a single equation in $\theta^{*'}$ by solving (A.6) for $S^{*'}$,

$$S^{*'}=\theta^{*'}+\sigma\Phi^{-1}\left(1-\frac{\frac{1}{2}-\Phi(\theta^{*'})}{F}\right), \tag{A.8}$$

and substituting it into (A.7). After some algebra, we obtain

$$\frac{\sigma_0}{\sqrt{\sigma_0^2+\sigma^2}}\Phi^{-1}\left(1-\frac{\frac{1}{2}-\Phi(\theta^{*'})}{F}\right)=\frac{\sigma}{\sigma_0\sqrt{\sigma_0^2+\sigma^2}}(\theta^{*'}-\theta_0)-\Phi^{-1}\left(\frac{w}{c+w}\right). \tag{A.9}$$

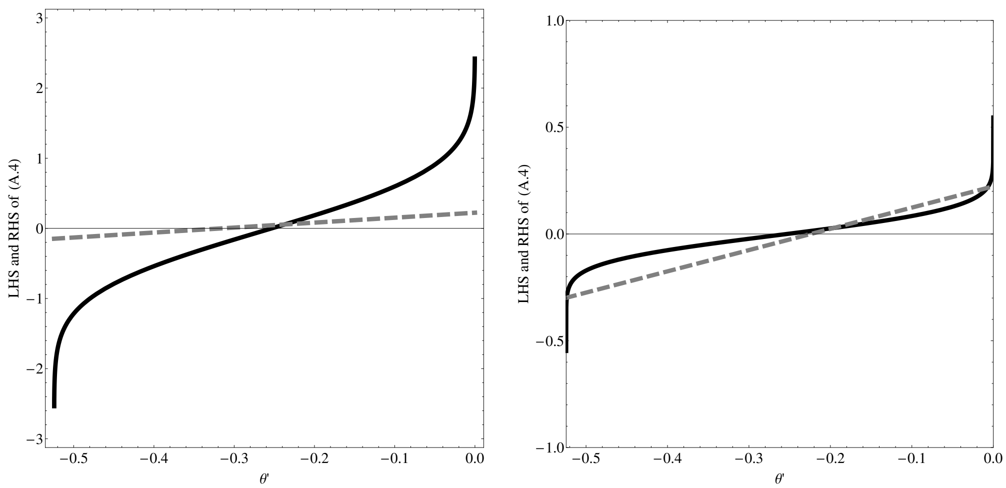


Figure A.4: The left-hand-side versus the right-hand-side of (A.9) for the case of a unique equilibrium (left) and a case that lacks uniqueness (right)

Equilibrium condition (A.9) has a unique solution for $\theta^{*'}$ as long as the signal S_i is sufficiently precise relative to the prior belief about θ. That is, as long as σ^2 is not too large relative to σ_0^2. Observe that the right-hand-side of (A.9) is linearly increasing in $\theta^{*'}$ with the slope $\frac{\sigma}{\sigma_0\sqrt{\sigma_0^2+\sigma^2}}$. Meanwhile, the left-hand-side of (A.9) is also increasing in $\theta^{*'}$ but it is non-linear (it is inverse S-shaped along the y-axis) and its slope may be smaller than $\frac{\sigma}{\sigma_0\sqrt{\sigma_0^2+\sigma^2}}$ for large values of σ^2 (relative to σ_0^2).[4] In these cases, (A.9) may no longer have a unique solution.

Figure A.4 illustrates the two cases. More formally, the partial derivative of the left-hand-side of (A.9) with respect to θ' is

$$\frac{\partial LHS}{\partial \theta'} = \frac{\sigma_0}{\sqrt{\sigma_0^2+\sigma^2}} \frac{e^{-\frac{\theta'^2}{2}+\text{erfc}^{-1}\left[2-\frac{1-2\Phi(\theta')}{F}\right]^2}}{F} > 0\,,$$

[4]The left-hand-side is defined only on the interval $\phi^{-1}(\frac{1}{2}-F) < \theta^{'} < 0$, with $\lim_{\theta'\to 0} = \infty$ and $\lim_{\theta'\to\phi^{-1}(\frac{1}{2}-F)} = -\infty$.

where $\text{erfc}^{-1}(x)$ is the inverse complementary error function (see e.g. **?**, 160),

$$\text{erfc}^{-1}(x) = 1 - \text{erf}(x) \quad \text{and} \quad \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt.$$

In turn, the slope of the left-hand-side of A.4 is steeper than the slope of its right-hand-side as long as

$$\frac{\sigma}{\sigma_0} < \min_{\theta'} \frac{e^{-\frac{\theta'^2}{2} + \text{erfc}^{-1}\left[2 - \frac{1-2\Phi(\theta')}{F}\right]^2}}{F}\,.$$

The above is a sufficient condition for a unique solution to (A.9). For the value of F used in the paper ($F = \frac{2}{10}$), for instance, this condition implies that uniqueness is guaranteed for any value of θ' as long as σ is at most 4.84 times as large as σ_0. (This uniqueness condition is satisfied for any realistic scenario as σ_0 is greater than σ when these are understood to be the standard deviation of the pre-election expectation and the election-day signal, respectively.)

The comparative statics of S^* are θ^* with respect to F, w, and c remain identical to the Uniform model. Differentiating (A.9) with respect to w, the left-hand-side is zero (since it does not depend on w) while the right-hand-side is negative as

$$\frac{\partial RHS}{\partial w} = -\frac{\sqrt{2\pi}\, c\, e^{\text{erfc}^{-1}\left[\frac{2w}{c+w}\right]^2}}{(c+w)^2} < 0\,.$$

This implies that an increase in w shifts the right-hand-side (A.9) downward while the left-hand-side is unchanged, resulting in a decrease in $\theta^{*'}$ (since the left-hand-side is increasing in θ'.) A similar argument confirms that $\theta^{*'}$ increasing in c, since

$$\frac{\partial RHS}{\partial c} = \frac{\sqrt{2\pi}\, w\, e^{\text{erfc}^{-1}\left[\frac{2w}{c+w}\right]^2}}{(c+w)^2} > 0\,.$$

Meanwhile, $\theta^{*'}$ is decreasing in F, as the partial derivative of the right-hand-side with respect to F is zero (since it does not depend on F) while the derivative of the left-hand-side is positive as

$$\frac{\partial LHS}{\partial F} = -\frac{\sigma_0 \sqrt{\frac{\pi}{2}}\, \text{erf}\left[\frac{\theta'}{\sqrt{2}}\right] \; e^{\text{erfc}^{-1}\left[2+\frac{\text{erf}\left[\frac{\theta'}{\sqrt{2}}\right]}{F}\right]^2}}{F^2\sqrt{\sigma_0^2+\sigma^2}} > 0\,,$$

where $\text{erf}\left[\frac{\theta'}{\sqrt{2}}\right] < 0$. The same relationship with respect to F, w, and c holds for S^* since $S^{*'}$ is increasing in $\theta^{*'}$ according to (A.8).

Finally, consider the incumbent's optimal choice of the reward factor w in light of the pre-election belief θ_0 about his popularity. The equivalent of the incumbent's expected payoff in (7) is

$$\Pr\left[\theta \geq \theta^*\right] b - wE[R] = \left[1 - \Phi\left(\frac{\theta^{*'} - \theta_0^{'}}{\sigma_0}\right)\right] b - w\int_{-\infty}^{\infty} E[R|\theta^{'}]f(\theta^{'})d\theta^{'}. \qquad \text{(A.10)}$$

In (A.10), equilibrium probability of the incumbent's victory $\Pr\left[\theta \geq \theta^*\right]$ corresponds to one minus the cumulative distribution function of the $\mathcal{N}(\theta_0^{'}, \sigma_0^2)$ distribution evaluated at $\theta^{*'}$, and the incumbent's expected national-level election outcome $E[R|\theta^{'}]$ is evaluated with respect to the distribution of $\theta^{'}$; $f(\theta^{'})$ thus corresponds to the probability density function of the $\mathcal{N}(\theta_0^{'}, \sigma_0^2)$ distribution. In turn,

$$\int_{-\infty}^{\infty} E[R|\theta^{'}]f(\theta^{'})d\theta^{'} = \int_{-\infty}^{\infty}\left(E[S_i|\theta^{'}] + F\phi(\theta^{'})\right) f(\theta^{'})d\theta^{'}. \qquad \text{(A.11)}$$

In this expression, $E[S_i|\theta^{'}]$ is the incumbent's expected genuine national-level vote-share when his popularity is $\theta^{'}$,

$$E[S_i|\theta^{'}] = \int_{-\infty}^{\infty} S_i f(S_i^{'}|\theta^{'})\, dS_i^{'} = \int_{-\infty}^{\infty} \Phi(S^{'}) f(S_i^{'}|\theta^{'})\, dS_i^{'},$$

which is evaluated with respect to the distribution of $S_i^{'}$; $f(S_i^{'}|\theta^{'})$ thus corresponds to the probability density function of the $\mathcal{N}(\theta^{'}, \sigma^2)$ distribution. Meanwhile, $\phi(\theta^{'})$ in (A.10) is the fraction of agents that engage in fraud in equilibrium when the incumbent's popularity is $\theta^{'}$ and corresponds to one minus the cumulative distribution function of the $\mathcal{N}(\theta', \sigma^2)$ distribution evaluated at $S^{*'}$,

$$\phi = 1 - \Phi\left(\frac{S^{*'} - \theta'}{\sigma}\right).$$

In equilibrium, the incumbent maximizes the expected payoff in (A.10) with respect to w while treating θ^*, S^*, and ϕ as functions of w.

This optimization problem is too mathematically complex to be analytically tractable and w^* must be found numerically. The parameter values c=1, $F = \frac{2}{10}$, $b = 100$, $\sigma^2 = \frac{1}{100}$, $\theta_0 = \frac{1}{2}$ (which implies $\theta_0^{'} = 0$), and $\sigma_0^2 = 1$, for instance, result in $w^* = 4.48$, $\theta^* = 0.33$, $S^* = 0.30$, and $\phi^* = 0.83$. That is, agents engage in fraud only if the incumbent's popularity in their precinct is greater than 30% and fraud secures the incumbent's victory if his national-level popularity is greater than 33%, or equivalently, when at least 83% of agents participate in fraud. This example is illustrated in Figure A.5, which plots the effect of the incumbent's actual national-level popularity θ on the equilibrium and needed levels of fraud as a solid black line, assuming that the incumbent's actual popularity θ will turn out as expected (i.e. $\theta = \theta_0 = \frac{1}{2}$).[5]

The Normal model highlights particularly well the contrast between the rigidity of the equilibrium outcome when the incumbent's popularity is above θ^* and the resounding

[5]This is the analogue of Figure 3 from our earlier discussion.

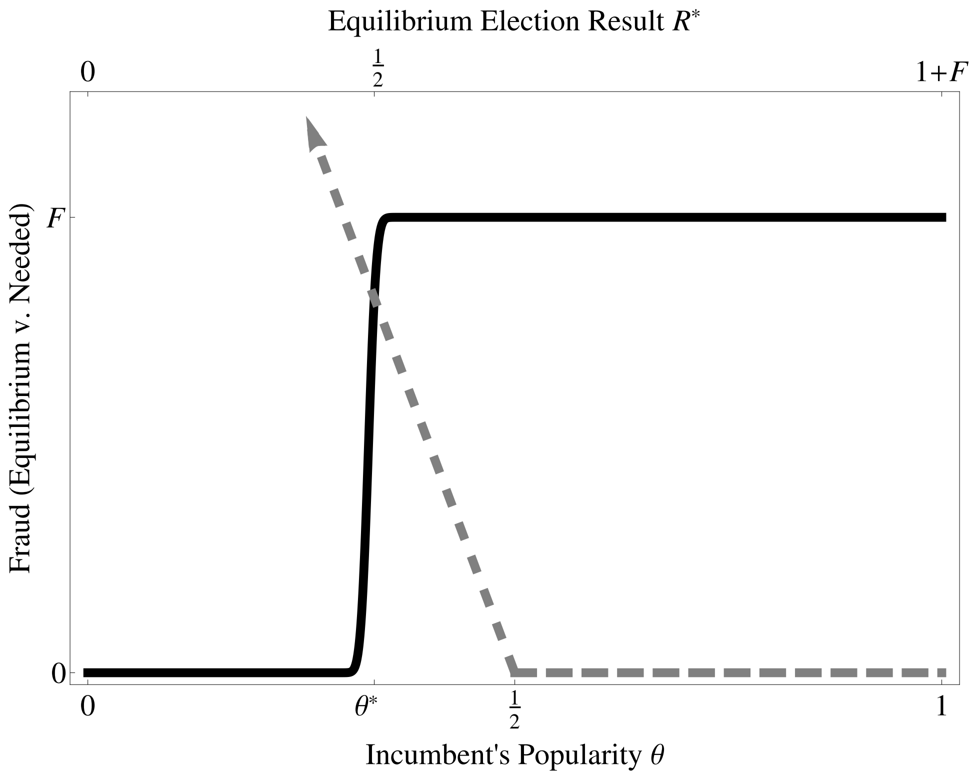


Figure A.5: The effect of the incumbent's actual national-level popularity θ on the equilibrium v. needed level of fraud in the Normal model

defeats that occur as the incumbent's popularity crosses below θ^*. In our example, agents receive highly precise signals, $\sigma^2 = \frac{1}{100}$. We see that when each agent has nearly perfect information about the incumbent's national-level popularity θ, shifts in the agents' perception of the incumbent's popularity result in herd-like coordination: On the one hand, virtually all agents conduct fraud on behalf of the incumbent regardless of the actual value of θ as long as $\theta > \theta^*$; on the other hand, a minor shift in the incumbent's popularity from just above to just below θ^* results in his defeat by a margin of about $F\%$ as virtually all agents change their behavior from conducting fraud to refraining from it.

B.1 The 2012 Presidential Election: Multiples of Five

Likelihood ratio independence tests: In order to examine the over-representation of 0's and 5's, we round each candidate's vote share to the nearest multiple of 0.5, extract the unit and the first decimal place digits, and pool them into the twenty resulting digit pairs. Figure A.6 displays the distribution of these digit pairs and demonstrates that the multiples of five are indeed over-represented. Assuming that neighboring digits should be distributed approximately uniformly, we compute the G^2 statistic (**?**, 36) for the frequencies of 0.0 and 5.0 and the two digit pairs to their left and right. These are the digit pairs $\{9.0, 9.5, 0.0, 0.5, 1.0\}$ and $\{4.0, 4.5, 5.0, 5.5, 6.0\}$, respectively. The G^2 statistics (75.0 and 38.3 with $df = 4$) strongly suggest that these digit frequencies are not uniform (both p-values $= 0$). Once we exclude the digit pairs 0.0 and 5.0, however, the remaining digit frequencies are consistent with uniformity ($G^2 = 2.6$ and 2.8 with $df = 3$ implying p-values of 0.46 and 0.42, respectively). An analysis based on standardized residuals implies the same conclusion.

The perturbation approach: In order to construct the null distribution for digit frequencies, we perturb precinct-level turnout and vote shares as follows: We draw turnout τ_i from the binomial distribution $Binomial(N_i, T_i)$, where N_i and T_i are the number of registered voters and actual turnout in precinct i. We then draw vote shares $(\nu_i^1, \ldots, \nu_i^5)$ for candidates 1 through 5 from the multinomial distribution $Multinomial(\tau_i, (V_i^1, \ldots, V_i^5))$, where $(V_i^1, \ldots, V_i^5)$ are the actual vote totals for candidates 1 through 5 in precinct i.

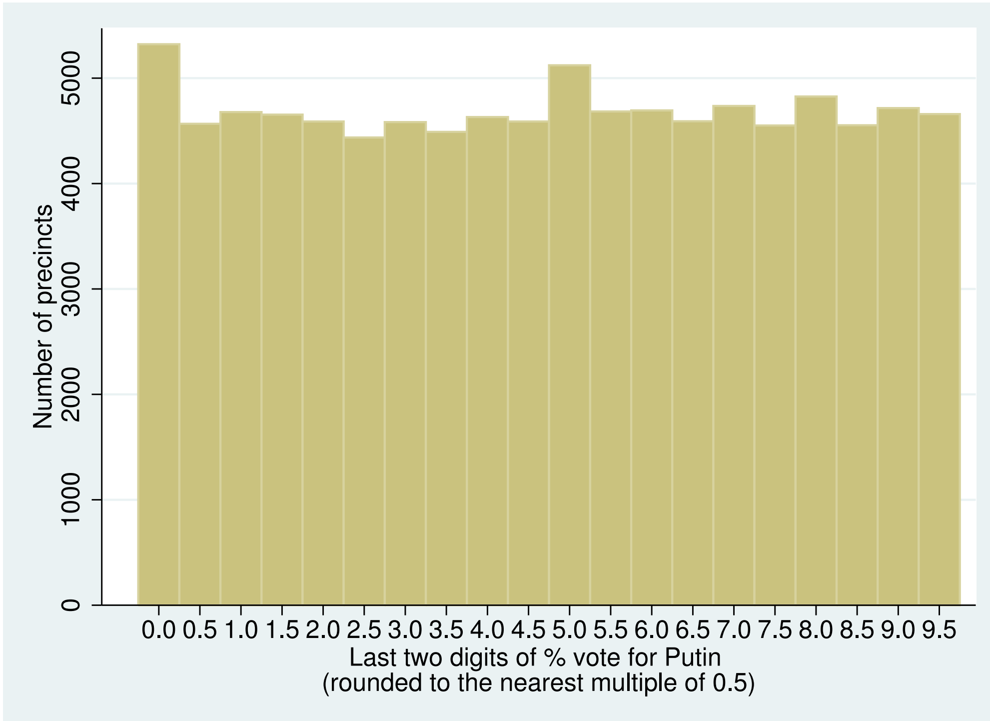


Figure A.6: The distribution of the pooled unit and the first decimal place digits in Putin's precinct-level vote share (after rounding to the nearest multiple of 0.5)

B.2 The 2012 Presidential Election: Alternative Kernel Density Estimate Bandwidths

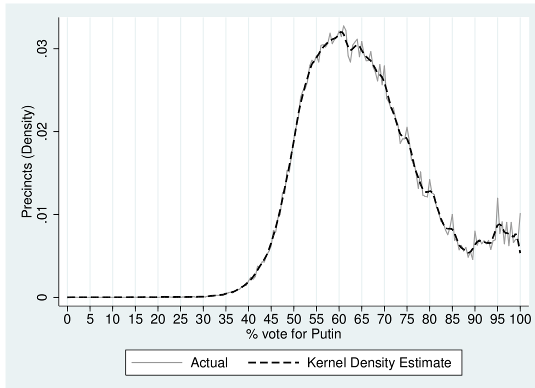
Following our theoretical analysis, we continue evaluating the prediction that the extent of fraud across individual precincts should be increasing in Putin's vote share. In order to evaluate this prediction, we develop a measure of ruggedness in the distribution of Putin's precinct-level results. In the paper, we report the difference between the empirical distribution of each candidate's precinct-level results and its optimal kernel density estimate.[6]

Due to space constraints, we were not able to report half and twice the optimal bandwidths as recommended by ?. We report these robustness checks in Figures A.7 and

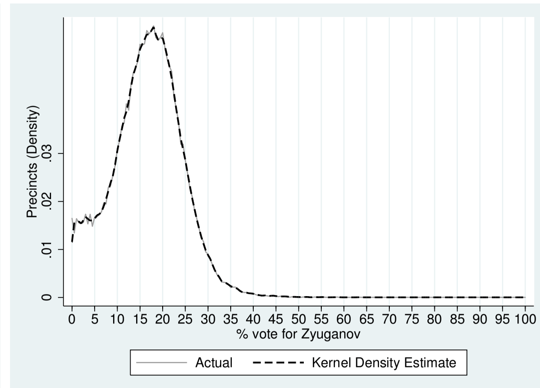
[6] The optimal bandwidth minimizes the mean integrated squared error based on a Gaussian kernel, and is 1.2 for Putin, 0.67 for Zyuganov, 0.19 for Mironov, 0.42 for Prokhorov, and 0.27 for Zhirinovsky.

A.8. The bandwidths do not change the conclusions from our previous analysis: the distribution of Putin's precinct-level results is still the only one in which departures from smoothness both coincide with the multiples of five and increase in his vote-share.

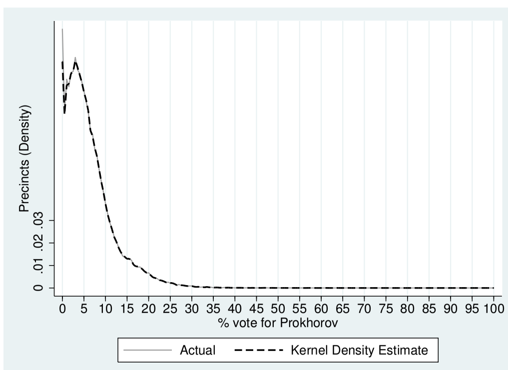
Figure A.9 presents the residuals from the kernel density estimate for Putin (diamonds), Zyuganov (squares), and the remaining three minor candidates (Prokhorov, Zhirinovsky, and Mironov). As in the main text, the 95th and 99th percentiles are based on the pooled residuals of all candidates but Putin. Once again, except for a few residuals clustered around 0 and 100 per cent, all residuals above the 95% and 99% percentiles belong to either Zyuganov or Putin. More importantly, as predicted by our theoretical model, the ruggedness of the distribution of precinct-level results is increasing Putin's and only Putin's vote-share and it is not sensitive to our choice of bandwidth.



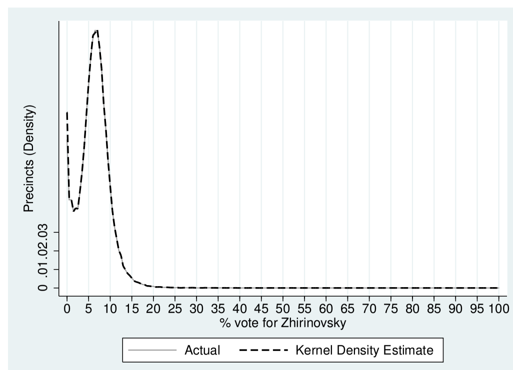
(a) Putin



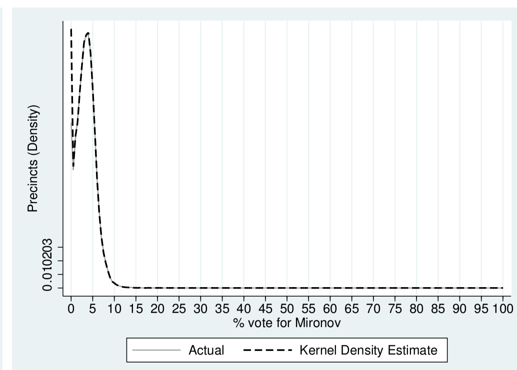
(b) Zyuganov



(c) Prokhorov

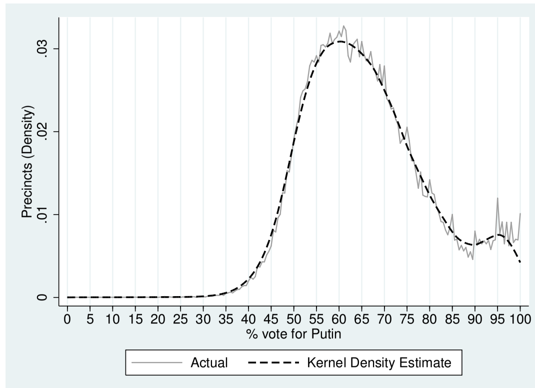


(d) Zhirinovsky

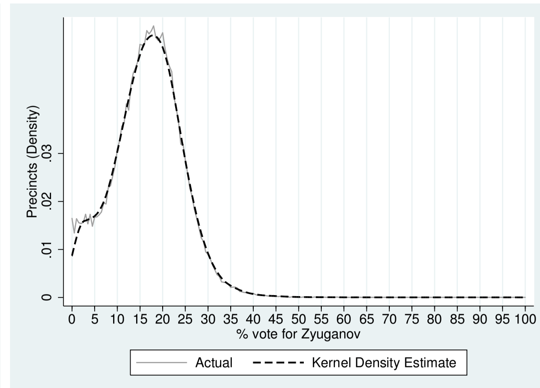


(e) Mironov

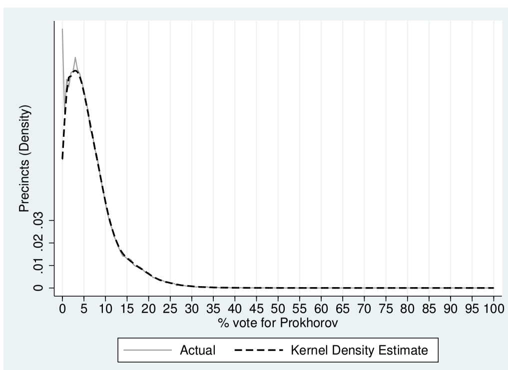
Figure A.7: The distribution (gray solid line) and kernel density estimate (black dashed line) of each candidate's precinct-level results using half the optimal bandwidth



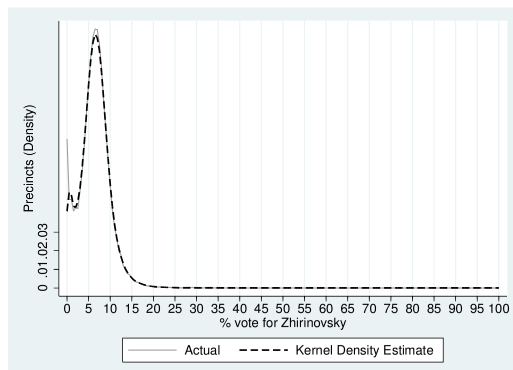
(a) Putin



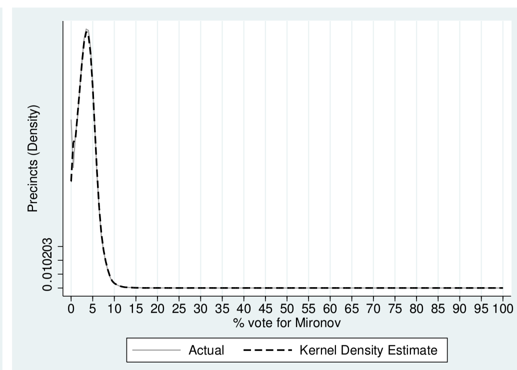
(b) Zyuganov



(c) Prokhorov

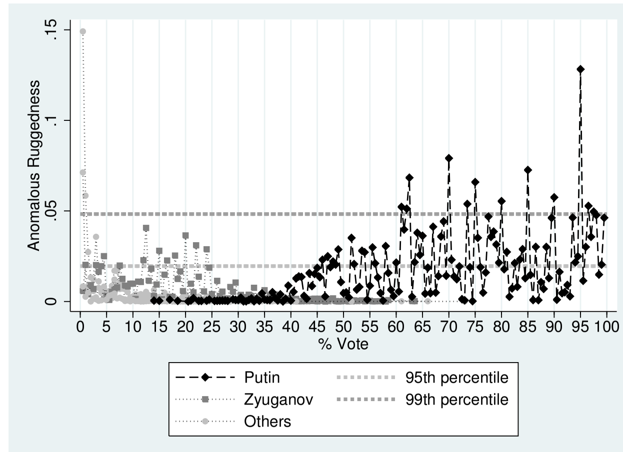


(d) Zhirinovsky

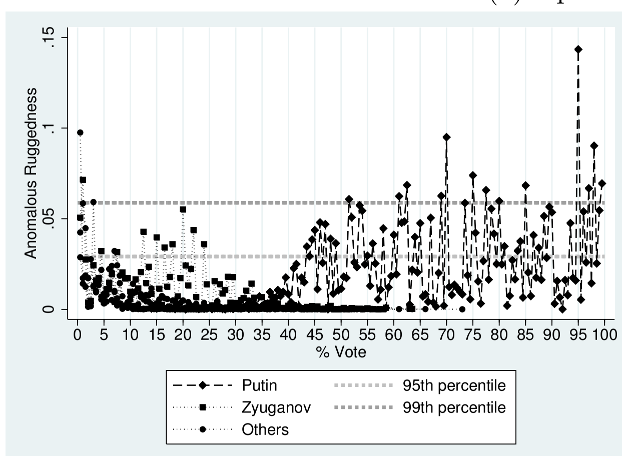


(e) Mironov

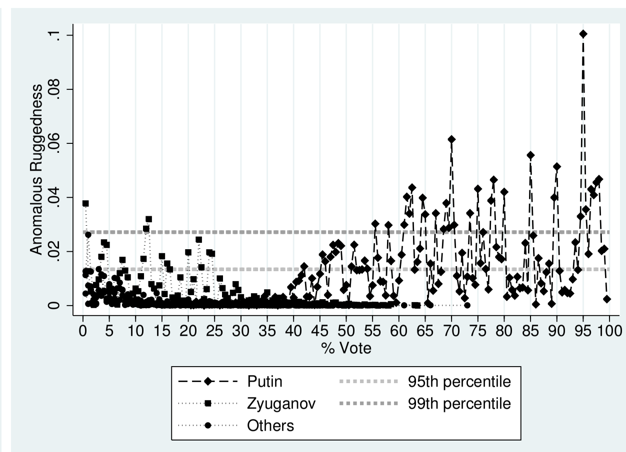
Figure A.8: The distribution (gray solid line) and kernel density estimate (black dashed line) of each candidate's precinct-level results using twice the optimal bandwidth



(a) Optimal bandwidth



(b) Twice the optimal bandwidth



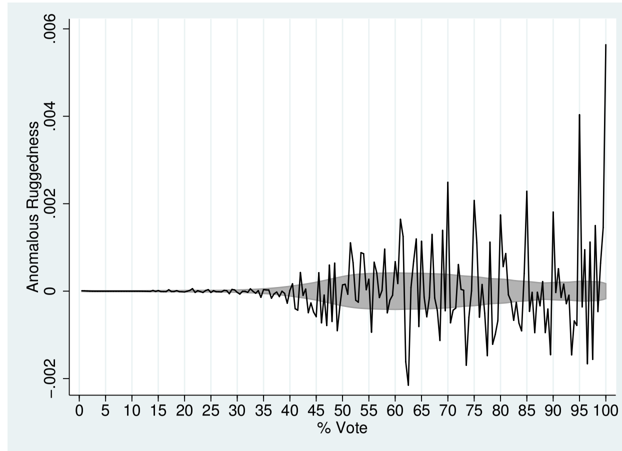
(c) Half the optimal bandwidth

Figure A.9: The distribution (gray solid line) and kernel density estimate (black dashed line) of each party's precinct-level results

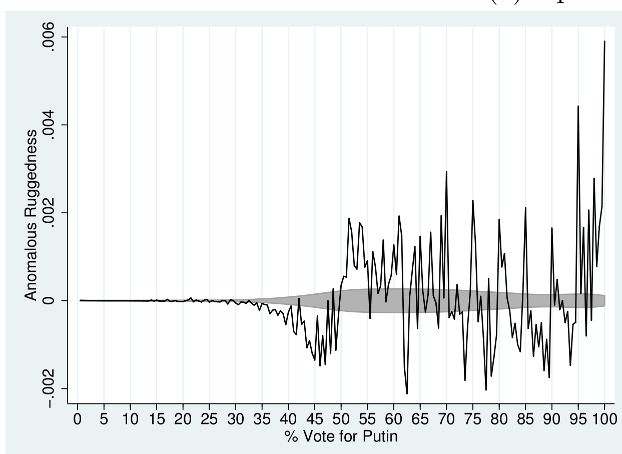
B.3 Asymptotic Confidence Intervals

In addition to the empirical confidence intervals reported in the main text and in the preceding section, we employ an alternative, theoretical benchmark for evaluating the ruggedness of Putin's precinct-level results. We compute the 95% asymptotic confidence intervals for the kernel density estimate of each candidate's distribution of precinct-level results and treat the observations that lie outside these confidence intervals as anomalously rugged. Figure A.10 displays the 95% asymptotic confidence intervals from the kernel density estimate for Putin (gray area) with the deviations of Putin's empirical distribution from its kernel density estimate (black line).

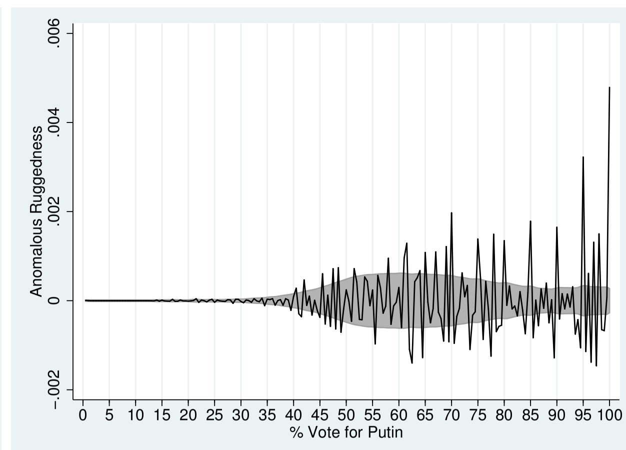
To test for robustness, we once again report the optimal, twice the optimal, and half the optimal bandwidths for our kernel density estimates. We judge residuals that fall outside of the 95% confidence interval as anomalously rugged, and report their absolute values in Figure A.11. Using these asymptotic confidence intervals, Figure A.12 plots the distribution of such residuals for all candidates. As previously explained, the 95th and 99th percentiles are based on the distribution of the pooled residuals of all candidates but Putin. We see that using this alternative, asymptotic benchmark, the ruggedness of the distribution of precinct-level results is again increasing Putin's and only Putin's vote-share – as anticipated by our theoretical arguments.



(a) Optimal bandwidth

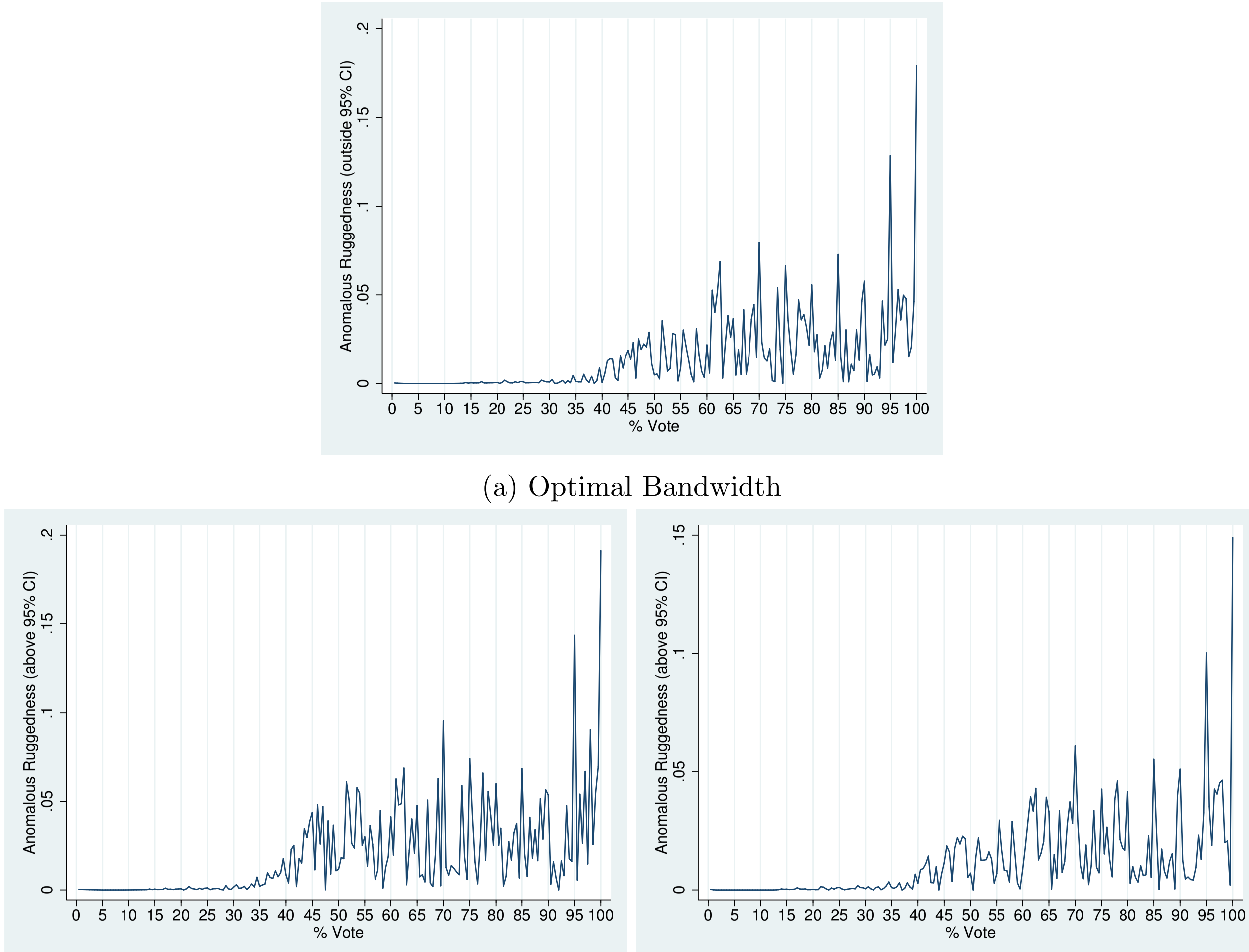


(b) Twice the optimal bandwidth



(c) Half the optimal bandwidth

Figure A.10: The 95% asymptotic confidence intervals (gray area) for the kernel density estimate's residuals (black line) of Putin's precinct-level vote share



(a) Optimal Bandwidth

(b) Twice the optimal bandwidth

(c) Half the optimal bandwidth

Figure A.11: The absolute value of the kernel density estimate residuals of Putin's precinct-level vote share that fall outside the 95% confidence interval

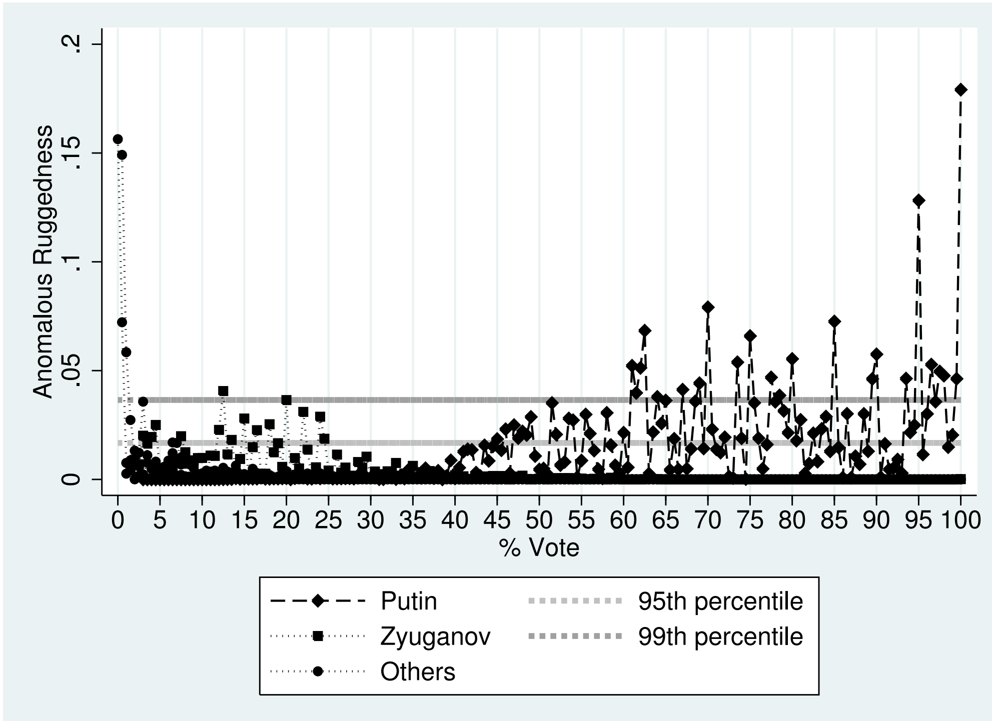
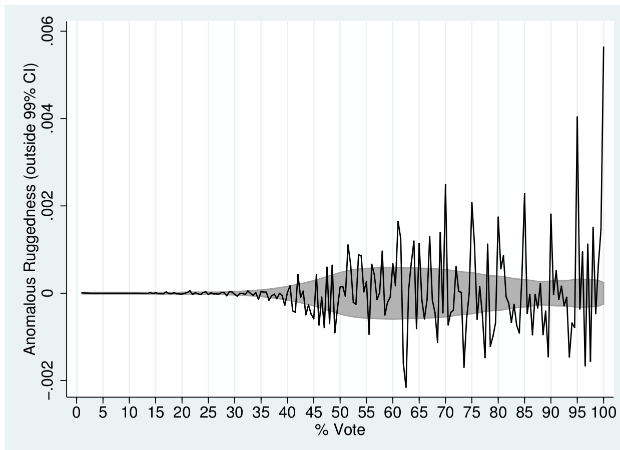
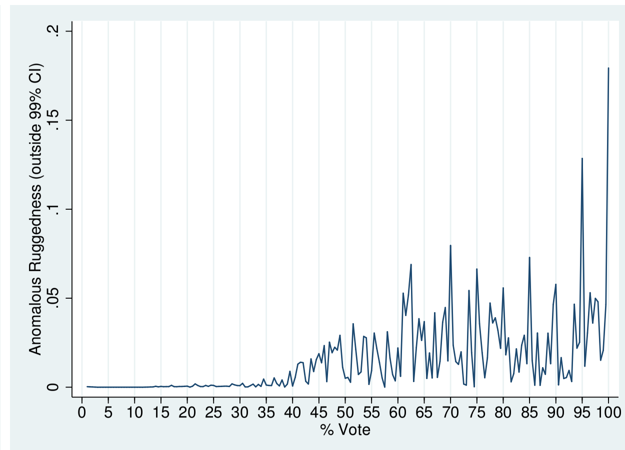


Figure A.12: The difference between the empirical distribution of each candidate's precinct-level results and the corresponding 95% confidence interval of the kernel density estimate

Finally, Figures A.13, A.14, and A.15 confirm our findings by reporting opposition candidate's 95% asymptotic confidence intervals, Putin's 99% asymptotic confidence intervals, and their absolute values.

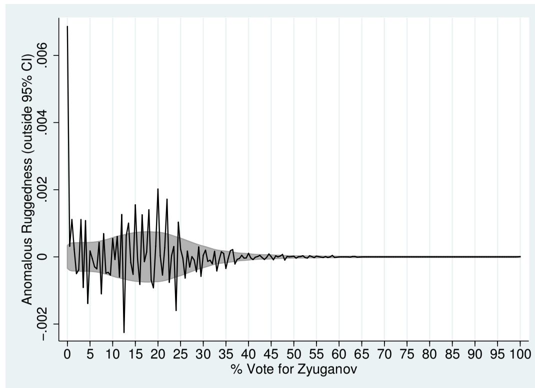


(a) 99% asymptotic confidence intervals (gray area) for the kernel density estimate's residuals (black line) of Putin's precinct-level vote share

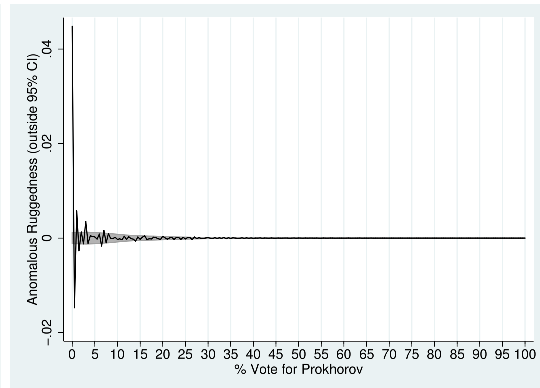


(b) The absolute value of the kernel density estimate residuals of Putin's precinct-level vote share that fall outside the 99% confidence interval

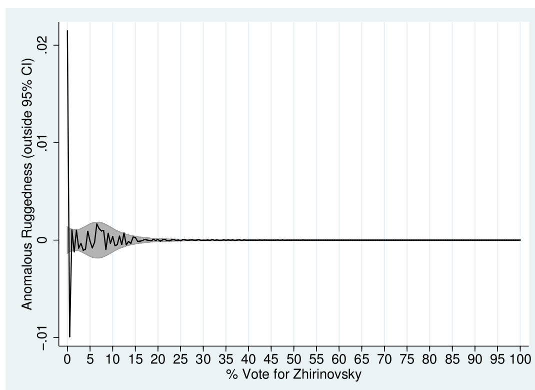
Figure A.13



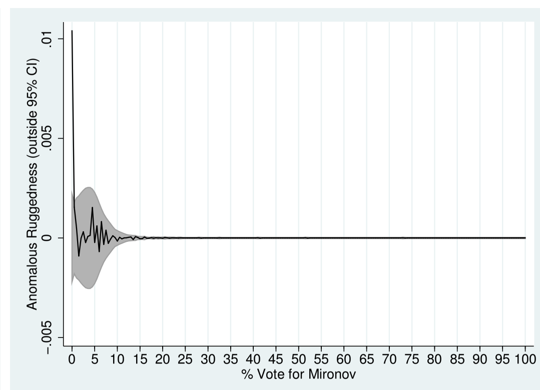
(a) Zyuganov



(b) Prokhorov

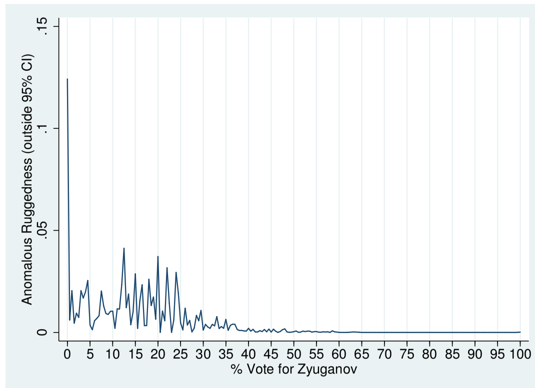


(c) Zhirinovsky

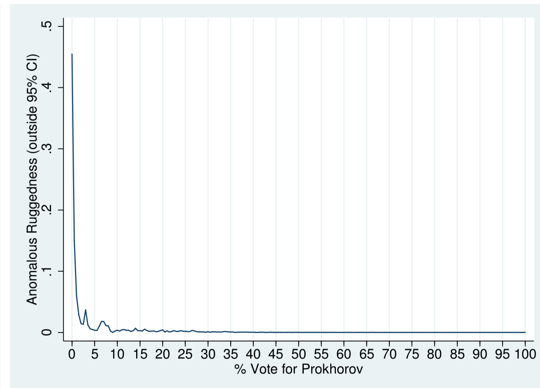


(d) Mironov

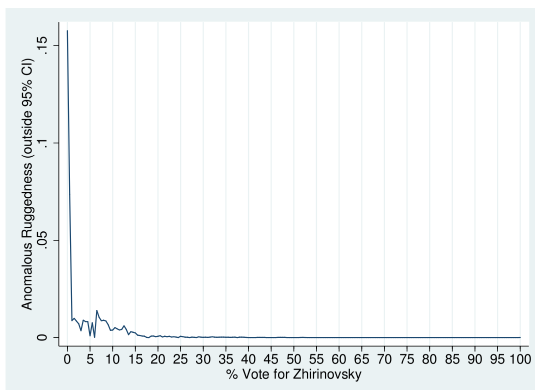
Figure A.14: The 95% asymptotic confidence intervals (gray area) for the kernel density estimate (black line) of each candidate's precinct-level vote share



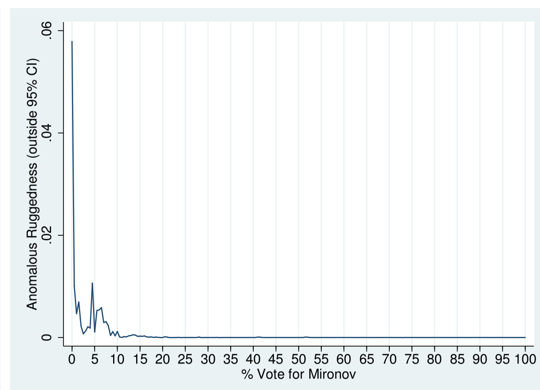
(a) Zyuganov



(b) Prokhorov



(c) Zhirinovsky



(d) Mironov

Figure A.15: The absolute value of the kernel density estimate residuals of each candidate's precinct-level vote share that fall outside the 99% confidence interval

B.4 Method of Fraud

Agents operating locally have two likely options for executing election fraud: the outright inflation of Putin's vote share by stuffing ballot boxes, or the stealing of votes from opposition candidates. Our analysis of the ruggedness in the distribution of precinct-level results in the 2012 Russian presidential election may already shed light on how fraud was conducted. Throughout our discussion above, Zyuganov was the only candidate other than Putin with a significant amount of ruggedness in his precinct-level results. Yet crucially, this ruggedness did not coincide with multiples of five and was not increasing in his vote share. This observation suggests that rather than stuffing ballots in order to round Putin's precinct-level vote share to some higher multiple of five, Putin's local operatives may have been instead stealing votes from Zyuganov. Stealing specifically from Zyuganov makes logistical sense: Zyuganov was the only major opposition candidate in this election and hence the only candidate with a number of votes large enough in most individual precincts that could be transferred to Putin's column in order to round his vote share to some higher multiple of five. In order to evaluate this first hypothesis, we add Putin's and Zyuganov's precinct-level votes and examine the ruggedness in the resulting distribution of vote shares. As Figure A.16 reveals, the significant ruggedness in the two candidates' individual vote-share distributions now disappears – supporting the hypothesis of vote-stealing from Zyuganov.

In order to test this hypothesis more thoroughly, we continue with our measure of ruggedness and compute the difference between the empirical distribution of each candidate's precinct-level results and its optimal kernel density estimate. Here, however, we use Putin and Zyuganov's combined distribution. Figure A.17 shows the difference between the empirical distribution of each candidate's precinct-level results and its kernel density estimate. We see that with the exception of a few spikes around 0 and 100 per cent (which

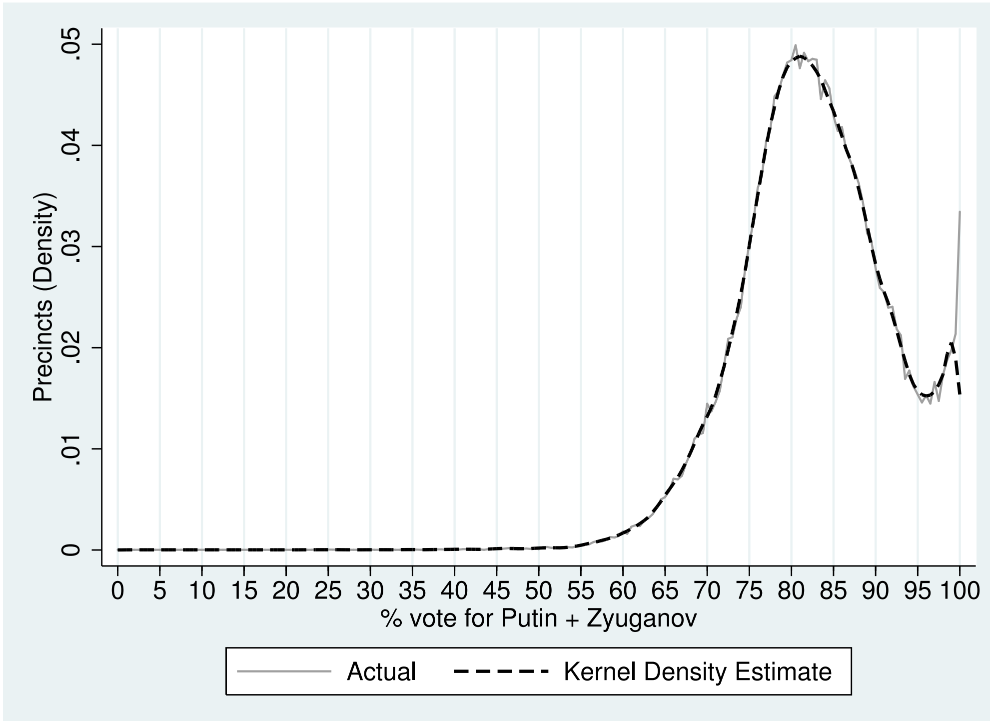


Figure A.16: The distribution (gray solid line) and kernel density estimate (black dashed line) of the sum of Putin's and Zyuganov's precinct-level vote share

are driven by the poor fit of the kernel density estimate for corner values), the ruggedness of Putin and Zyuganov's combined distribution is significantly smaller than the ruggedness of their individual distributions and close to the ruggedness of the remaining candidates.

The possibility remains, however, that these results are an artifact of the data; the smoothness of Putin and Zyuganov's distribution may only be a product of combining the two top-performing candidates rather than an indication that operatives specifically stole from Zyuganov. To explore this possibility, we take the candidate who received the least number of votes, Mironov, and inflate his vote share by 5.5, thereby approximating Zyuganov's average vote share across precincts. Since Mironov received only 3.9% of the total vote, he is the least-likely target for Putin's operatives. Figure A.18 reveals that when we combine Mironov's inflated vote share with Putin's vote share, the significant

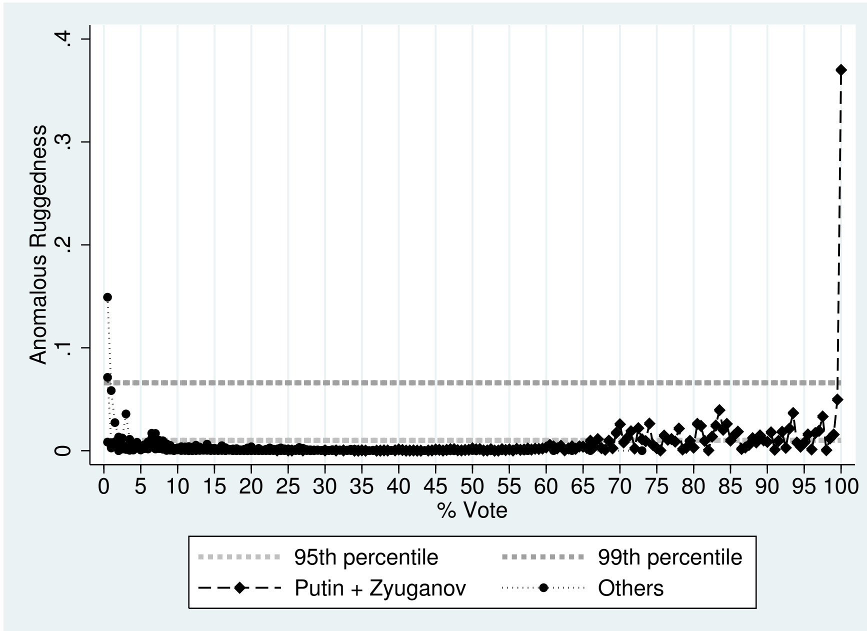


Figure A.17: The difference between the empirical distribution of each candidate's precinct-level results and its kernel density estimate

ruggedness in the two candidates' individual vote share distributions disappears. Because we are fairly certain that operatives did not steal from Mironov, this casts serious doubt on our evidence that operatives were stealing votes from Zyuganov.

To test for evidence of the alternative method of fraud, ballot box stuffing, we focus on turnout. If Putin's operatives were stealing from Zyuganov, we expect turnout to have no clear relationship with Putin's vote share (?). Figure A.19, however, reveals a strong relationship between turnout and Putin's vote share, suggesting that operatives did engage in ballot-box stuffing.

We find more evidence of this method when we plot the average percent turnout against Putin's vote share rounded to 0.5%. Figure A.20 reveals spikes in turnout around digits ending in 0.0 and 5.0, further highlighting the probability that operatives primarily used ballot box stuffing to round Putin's vote share to a higher multiple of 5.

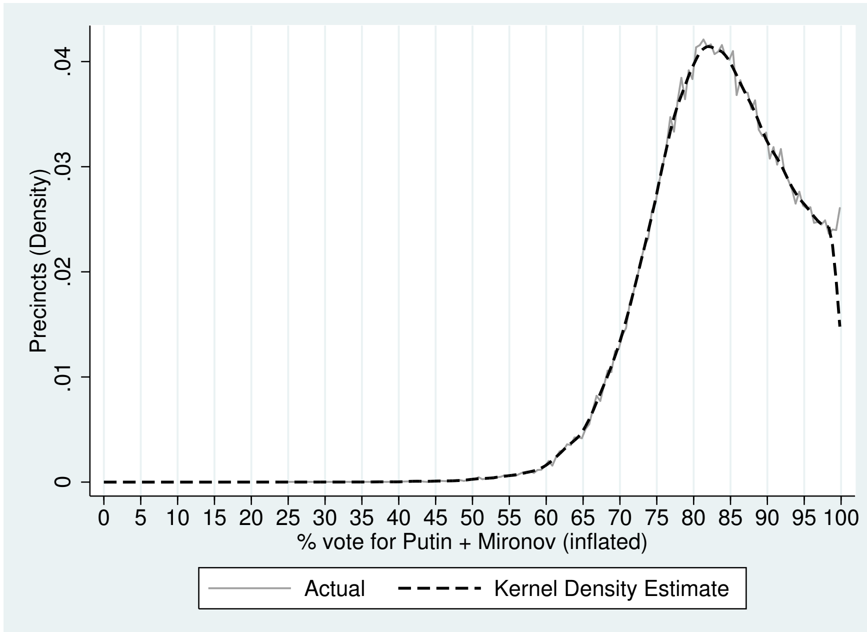


Figure A.18: The distribution (gray solid line) and kernel density estimate (black dashed line) of the sum of Mironov's (inflated) and Putin's precinct-level vote share

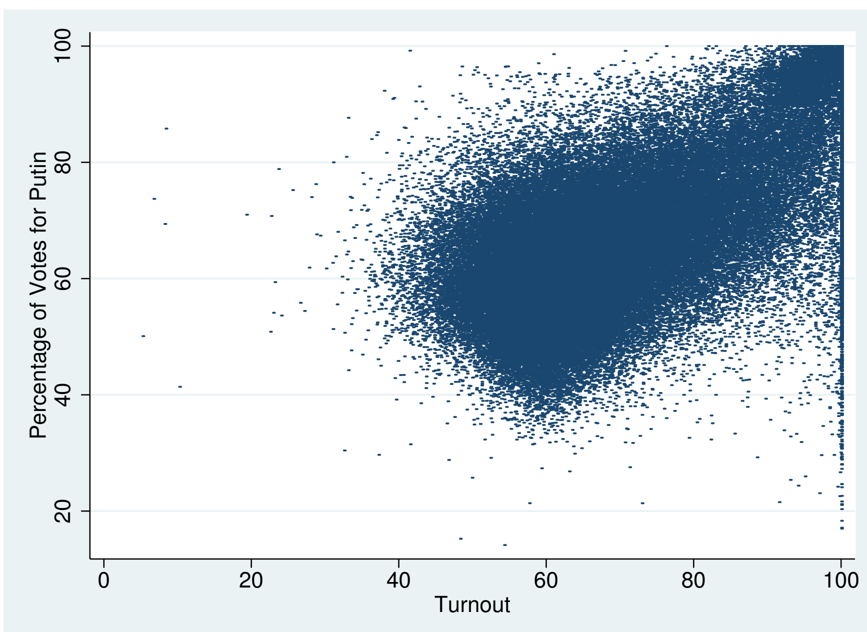


Figure A.19: Turnout and Putin's precinct-level vote share

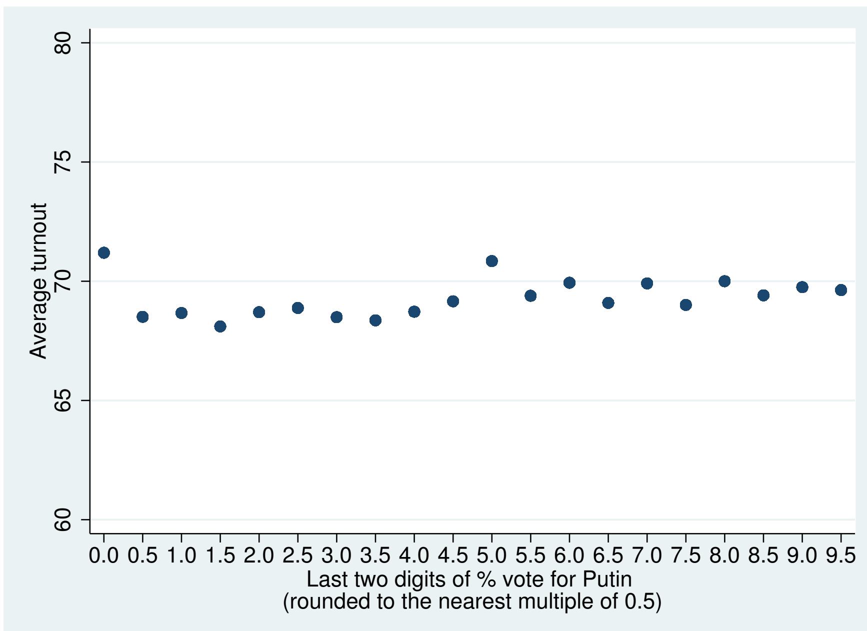


Figure A.20: Turnout and Putin's precinct-level vote share

B.5 Multiples of Five Analysis for Other Candidates

Although the empirical distributions of opposition candidates' vote shares do not appear to have a consistent pattern, this does not preclude the possibility that other distributions have a statistically significant over-representation of 5's and 0's. To explore this possibility, we perform the likelihood ratio independence tests for all opposition candidates' vote shares. Beginning with Putin's primary political opponent, Zyuganov's G^2 statistics around 0 and 5 are 180.5 and 9.4 respectively, with p-values of 0 and 0.052 respectively. Once 0 and 5 are removed, the G^2 statistics are 123.3 and 7.3, p-values = 0 and 0.064 respectively, indicating that while digit frequencies do not follow the expected uniform distribution, non-uniformity is not the result of an overabundance of 0's and 5's. Other candidates' likelihood ratio independence tests indicate the same conclusion. Mironov's G^2 statistic is 12000 and 2600 around 0 and 5 and 9500 and 2600 once the 0 and 5's are removed; Zhirinovsky's G^2 statistic is 1200 and 770.8 around 0 and 5, and 133.6 and 769.2 once the 0 and 5's are removed; Prokhorov's G^2 is 4400 and 109.4 around 0 and 5, and 2600 and 109.4 once the 0 and 5's are removed (all p-values = 0).

Because there are many precincts in which the three minor candidates, Mironov, Zhirinovsky, and Prokhorov, received close to 0% of the vote, this may explain the lack of uniformity and over-representation of zeros. Once we remove all precincts in which each candidate received less than 1% of the vote, Putin's G^2 statistics still strongly indicate an over-representation of 0's and 5's. (The G^2 statistics are 74.9 and 38.5 around 0 and 5, (p-values = 0), and 2.6 and 1.5 without 0 and 5's (p-values = 0.455 and 0.672).

Opposition candidates, however, still deviate from the expected uniform distribution of last digits. Zyuganov's G^2 statistic around 0 and 5 is 27.8 and 9.4(p-value = 0 and 0.052 respectively) and 26.8 and 7.3 without 0 and 5's (p-values = 0 and 0.064). Prokhorov's G^2 statistic around 0 and 5 is 764.9 and 109.4 and 677.9 and 109.4; Zhirinovsky's G^2 statistic is 842.7 and 772.3 around 0 and 5, and 808 and 770.7 without the 0's and 5's; Mironov's G^2 statistics are 5200 and 2600 around 0 and 5, and 4400 and 2600 without the 0's and 5's (p-values = 0), all of which still indicates that the lack of uniformity is unrelated to the over-representation of 0's and 5's.

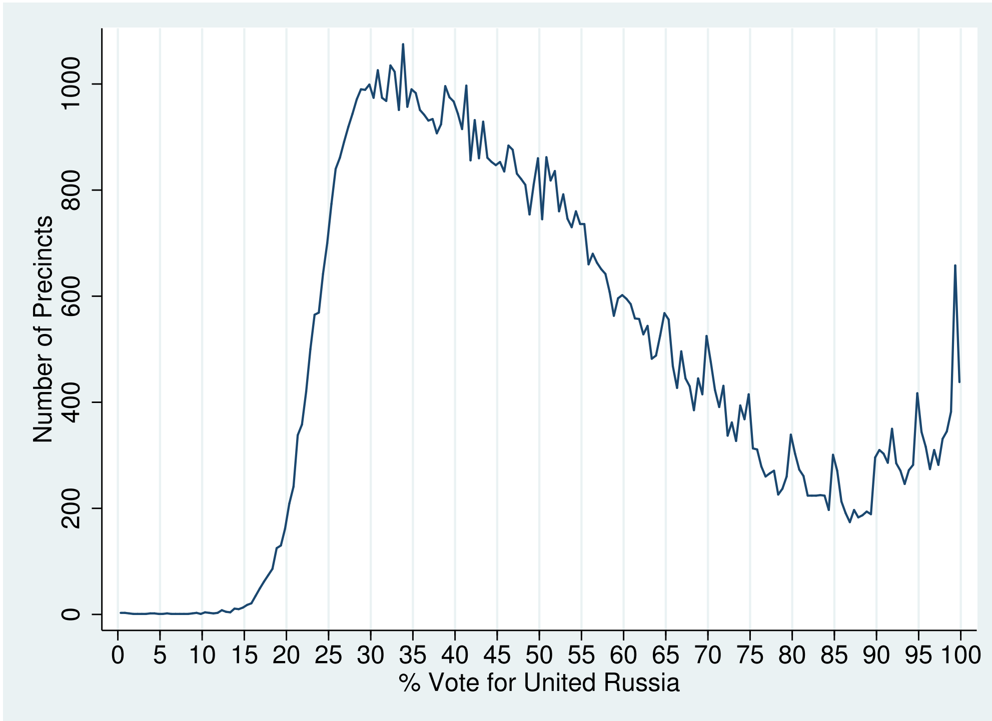


Figure A.21: The distribution of United Russia's precinct-level vote share in the 2011 parliamentary election

C.1 Results for 2011 Parliamentary Election

In our analysis of the 2011 legislative election in Russia, we take advantage of the same form of fraud evident in the 2012 presidential election: the rounding of the incumbent's precinct-level vote share to a higher multiple of five. Focusing on the hypothesis that the extent of fraud should be increasing in the incumbent's vote share, we first establish that 0's and 5's are over-represented in United Russia's vote share. Figure A.21 plots the distribution of United Russia's vote share across more than 90,000 precincts. Even more prominent in this election is the suspicious lack of smoothness coinciding with the multiples of five, especially between 65%-100%. In order to more rigorously examine the over-representation of 0's and 5's, we round each candidate's vote share to the nearest multiple of 0.5, extract the unit and the first decimal place digits, and pool them into the

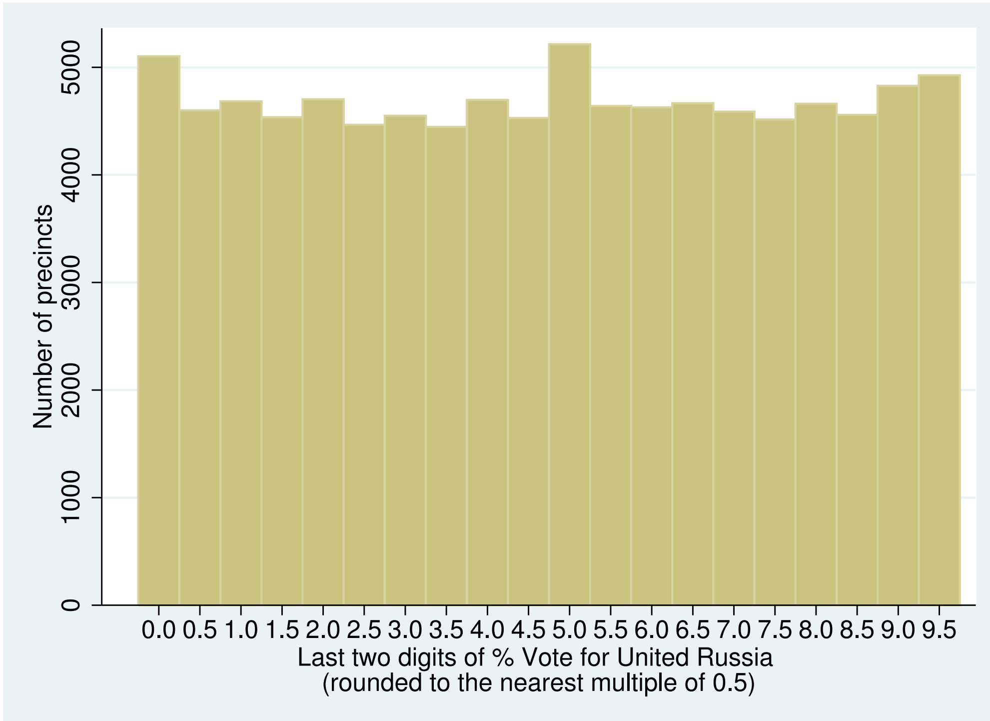
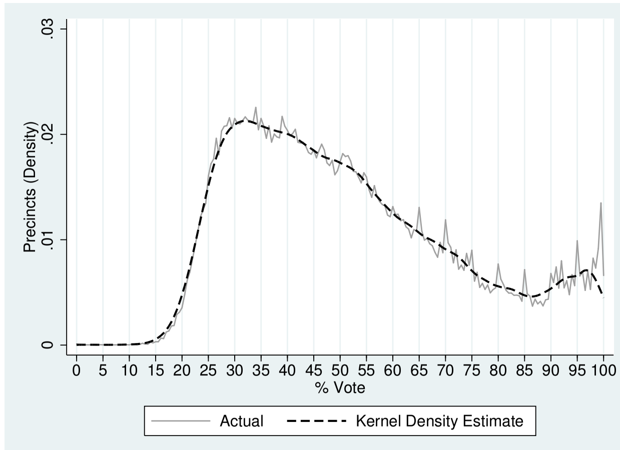


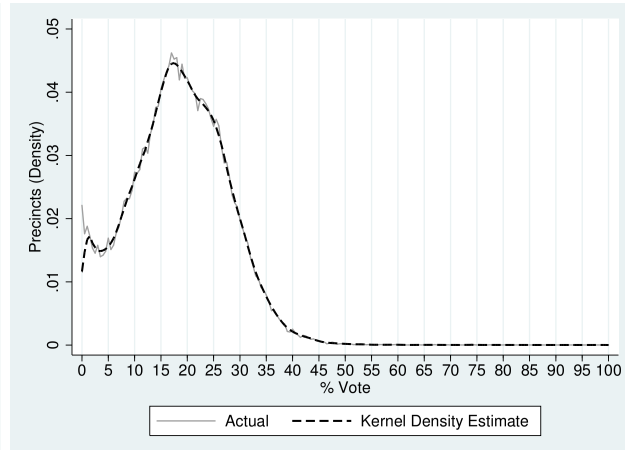
Figure A.22: The distribution of the pooled unit and the first decimal place digits in United Russia's precinct-level vote share (after rounding to the nearest multiple of 0.5)

twenty resulting digit pairs. Figure A.22 displays the distribution of these digit pairs and demonstrates that the multiples of five are indeed over-represented. Assuming that digits should be distributed uniformly, we compute the G^2 statistics for the frequencies of 0.0 and 5.0, and the two digit pairs to their left and right. The G^2 statistics (32.6 and 60.76 with $df = 4$), indicate that the digit frequencies are not uniform (both p-values = 0). Like the 2012 presidential election, the G^2 statistics excluding 5 suggest that departures from smoothness are the result of the over-representation of multiples of 5 (3.18, p=0.37). Unlikes the 2012 election, however, the G^2 statistics excluding 0.0 suggests that neighboring digits are not distributed uniformly (13.35, p=0.004). Departures from smoothness thus cannot be attributed solely to the over-representation of multiples of 5.

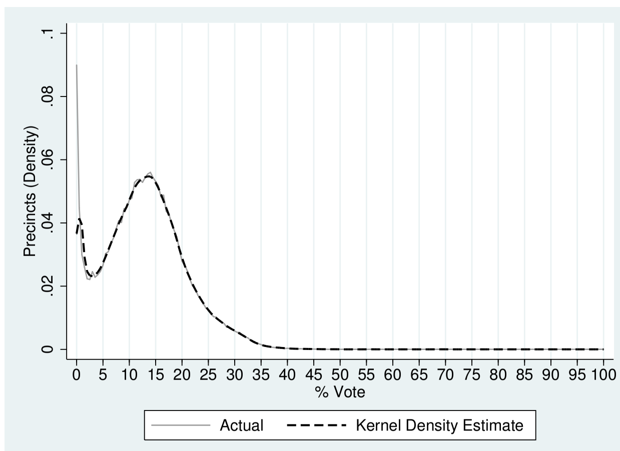
Turning to our primary empirical test, we assess whether departures from smoothness are increasing in United Russia's vote-share. We use the same measure developed in the empirical section of the main text by taking the difference between the empirical distribution of a candidate's precinct-level vote share and its optimal kernel density estimate. Figure A.23 plots the kernel density estimate for each party's precinct-level results by a black dashed line along with their actual empirical distribution (gray solid line). The empirical distributions for the two minor parties (A Just Russia and LDPR) almost exactly conform to their kernel density estimates, while the Communist Party shows a small amount of ruggedness unrelated to multiples of 5. Significant ruggedness corresponding with multiples of 5 is present only for United Russia, and appears to be increasing in United Russia's share of the vote.



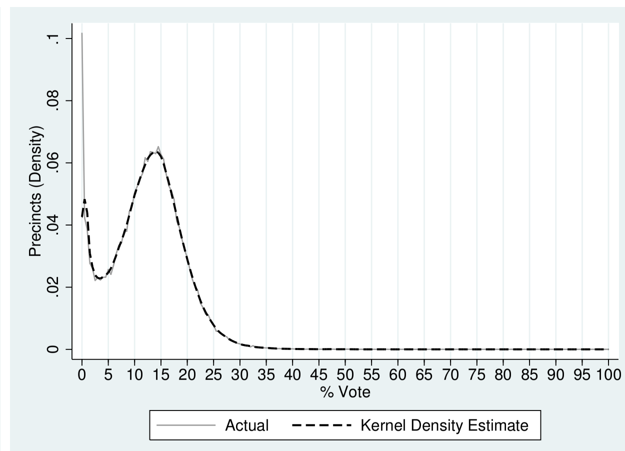
(a) United Russia



(b) Communist Party



(c) A Just Russia



(d) LDPR

Figure A.23: The distribution (gray solid line) and kernel density estimate (black dashed line) of each party's precinct-level results

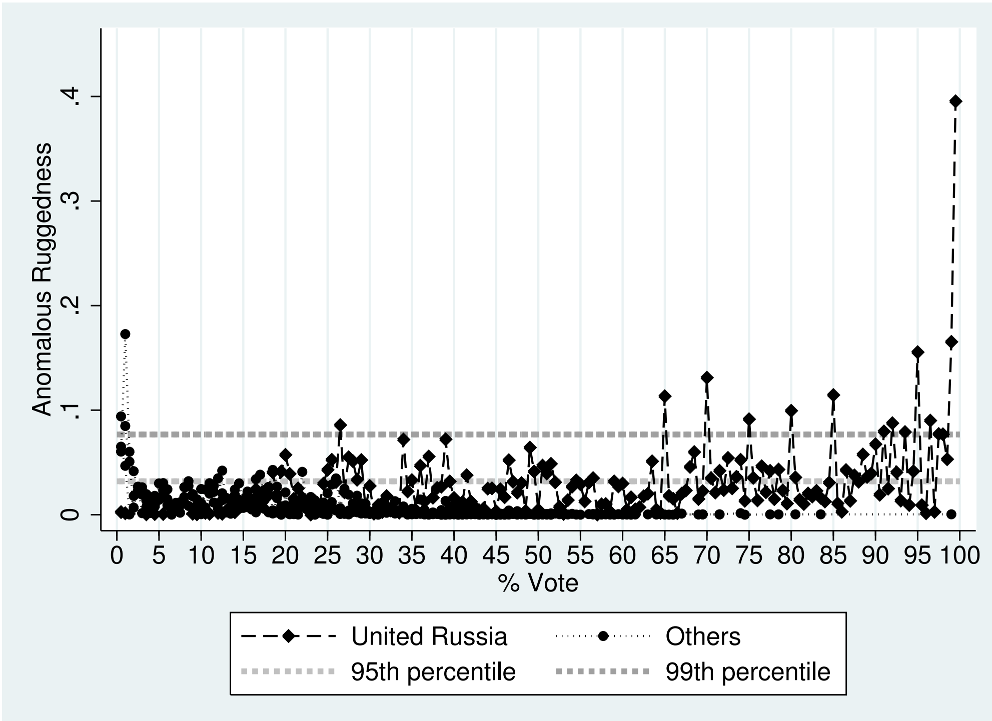


Figure A.24: The difference between the empirical distribution of each party's precinct-level results and its kernel density estimate

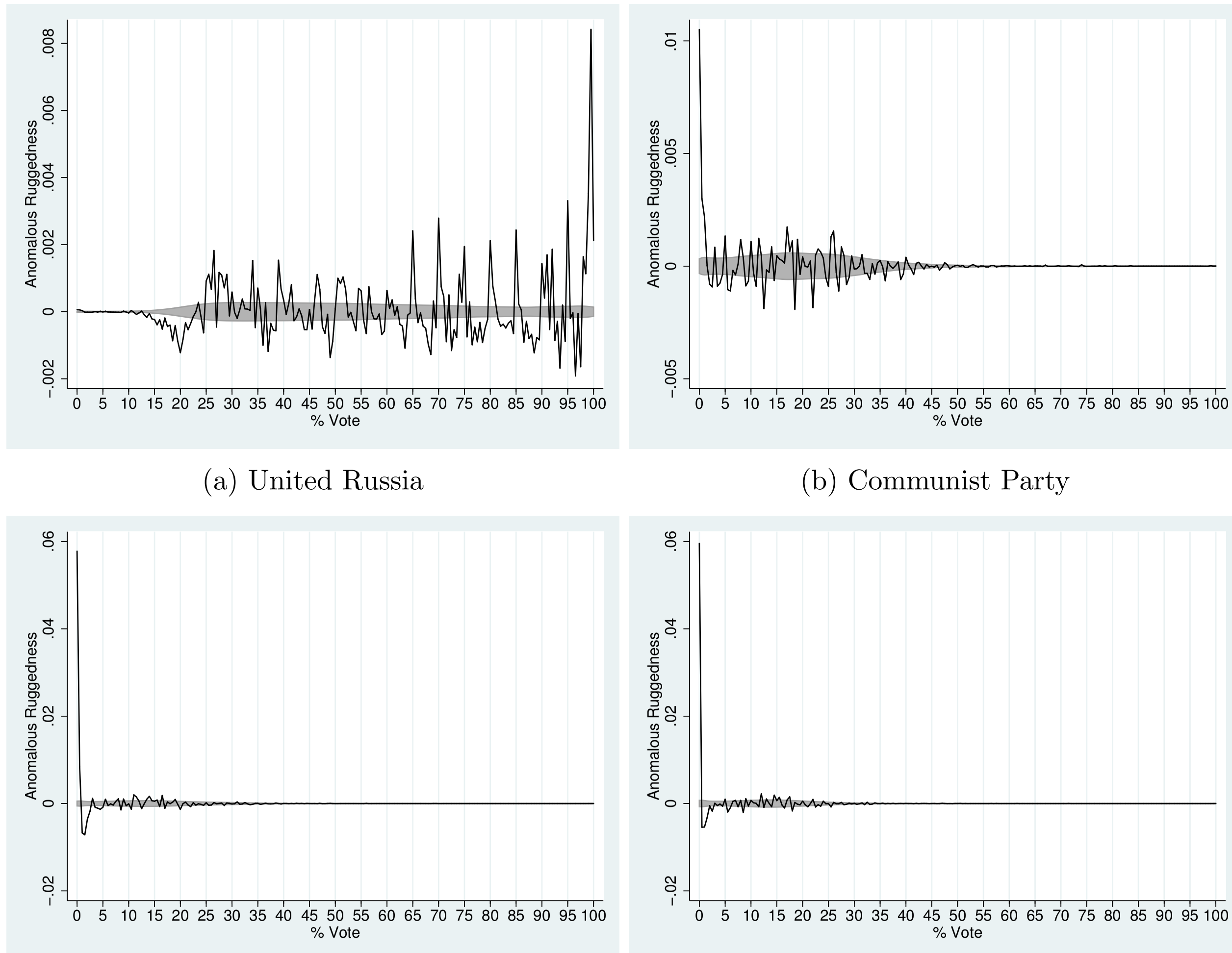
In order to quantify United Russia's distribution ruggedness, we use the same empirical and theoretical benchmarks presented in the main text. First, we judge United Russia's distribution by the standard of its three competitors. We first calculate the difference between the empirical distribution of United Russia's competitors' precinct-level results and their kernel density estimates, pool these residuals, and use their 95th and 99th percentiles as a benchmark for judging how anomalous Putin's ruggedness is. Figure A.24 compares United Russia's residuals (diamonds) to its competitors residuals, and the 95th and 99th percentiles. Consistent with our analysis of the 2012 presidential election, we use the optimal bandwidth for our kernel density estimate, and both twice and half the optimal bandwidth to check for robustness.[7] With the exception of a few residuals at the lower end

[7]The optimal bandwidth for United Russia is 1.89, for the LDPR is 0.62, for A Just Russia is 0.69, and for the Communist Party is 0.85.

of the distribution due to minor parties receiving 0 or almost 0% of the vote, United Russia's residuals are the only exceeding the 99 percentile. Additionally, while not as striking as the 2012 presidential election, United Russia's residuals are nevertheless the only residuals increasing in the party's vote share.[8]

Using an alternative theoretical benchmark, we compute the 95% and 99% asymptotic confidence intervals for the kernel density estimate of United Russia's results and treat the empirical observations outside of these confidence intervals as anomalously rugged. Figure A.25 demonstrates that once again, United Russia's residuals are significantly larger than those of the remaining party's residuals. Figure A.26 presents the absolute value of these residuals, making the relationship between United Russia's residuals and its vote share even more apparent. While the Communist Party's residuals are noticeably anomalous, they have no relationship with its vote share.

[8]When we regress United Russia's residuals on its vote share, the vote share coefficient is positive and statistically significant at the 0.01 significance level; the regression coefficient on the other parties' vote share is statistically significant but negative.



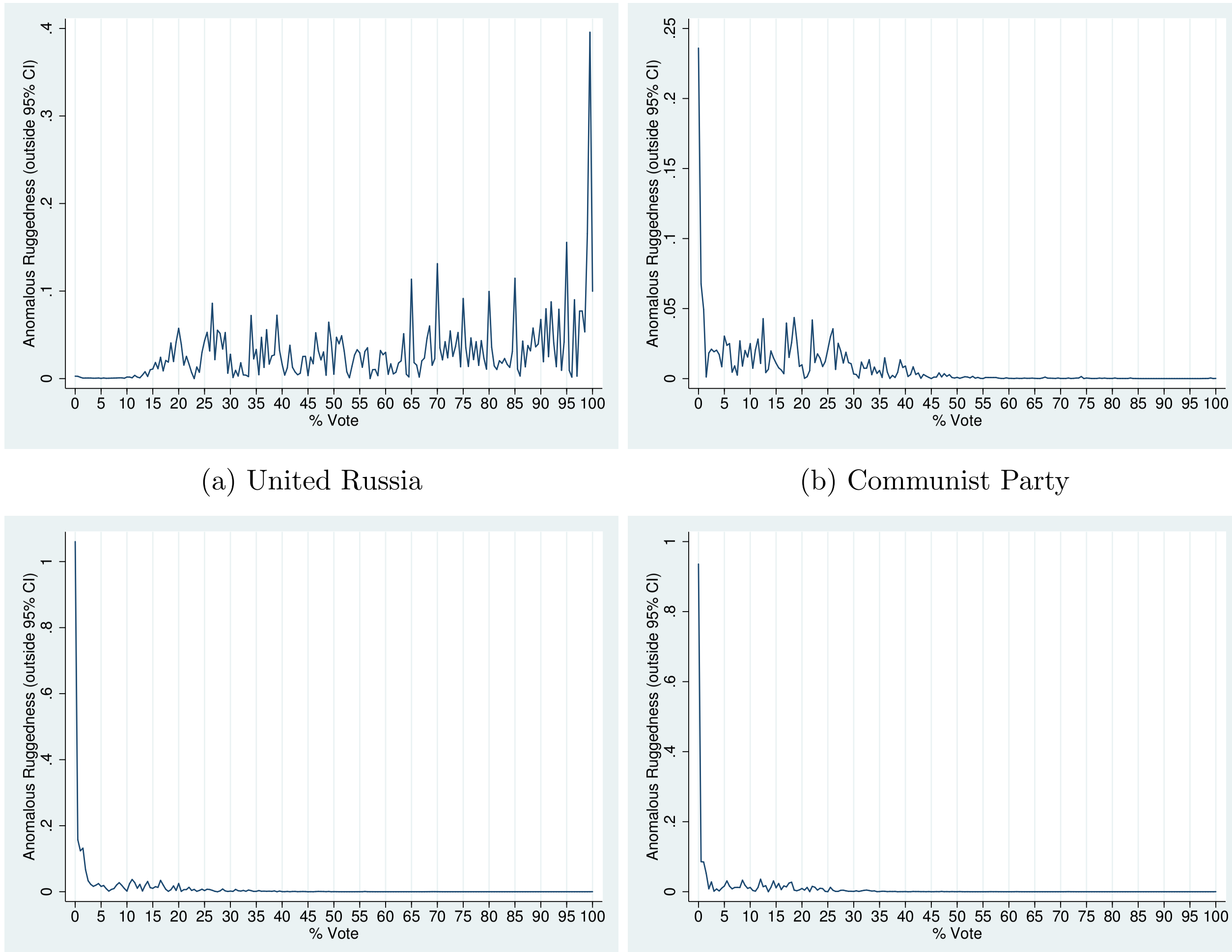
(a) United Russia

(b) Communist Party

(c) A Just Russia

(d) LDPR

Figure A.25: The 95% asymptotic confidence intervals (gray area) for the kernel density estimate (black line) of each party's precinct-level vote share



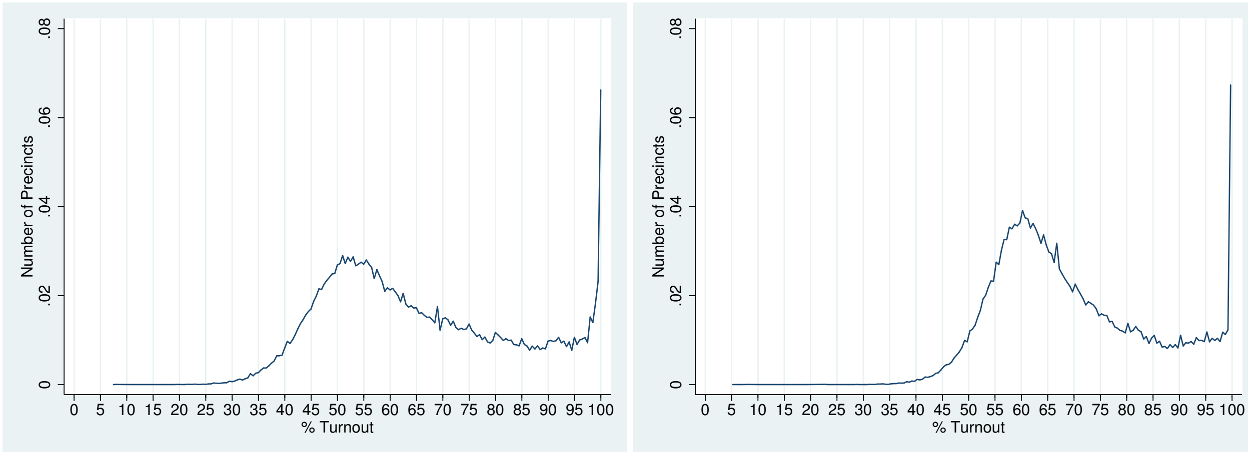
(a) United Russia

(b) Communist Party

(c) A Just Russia

(d) LDPR

Figure A.26: The absolute value of the difference between the empirical distribution of United Russia's precinct-level vote share and 95% confidence interval of its kernel density estimate



(a) 2011 parliamentary election

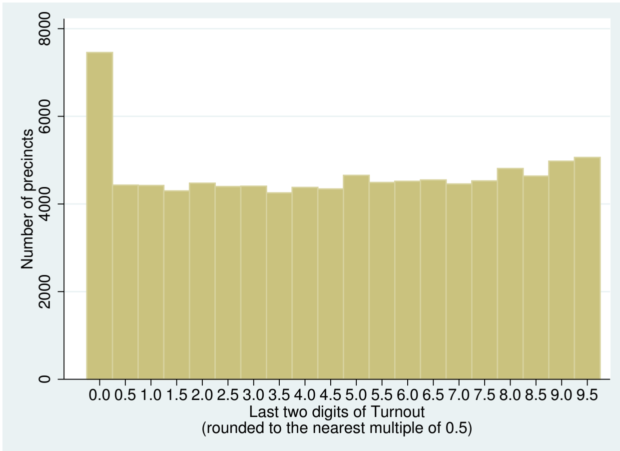
(b) 2012 presidential election

Figure A.27: The distribution of turnout in the 2011 parliamentary and 2012 presidential elections

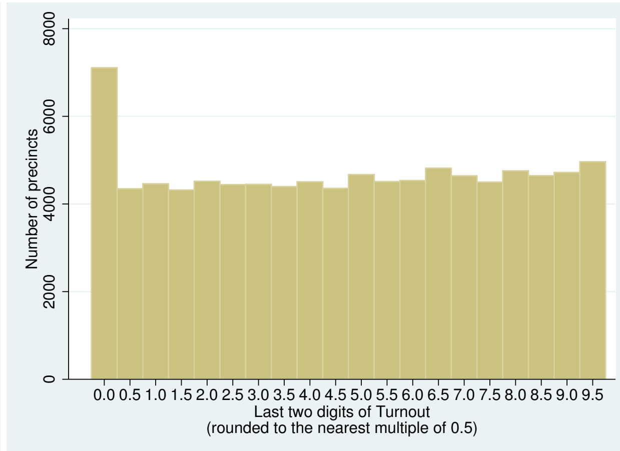
D.1 2011 Parliamentary and 2012 Presidential Elections: Analysis of Turnout

As a preliminary look at inflated turnout as a source of electoral fraud, Figure A.27 plots the distribution of turnout across the more than 90,000 precincts. Although the distribution display the familiar ruggedness present in Putin's vote share, the most noticeable outlier is the extreme bump in turnout at or near 100%. This, in conjunction with our analysis of the combined Putin and Zyuganov vote share density suggests that local operatives used two methods of electoral fraud: stealing from the second-place candidate, and if that did not offer enough votes to reach a satisfactory vote share, inflating turnout. As expected, Figure A.28 confirms the presence of an overabundance of 0's, largely driven by the large number of precincts reporting 100% turnout.

We repeat our main empirical test by assessing whether departures from smoothness are increasing in turnout. Figure A.29 plots the kernel density estimate for turnout by a black dashed line along with its actual empirical distribution (gray solid line). Turnout is



(a) 2011 parliamentary election



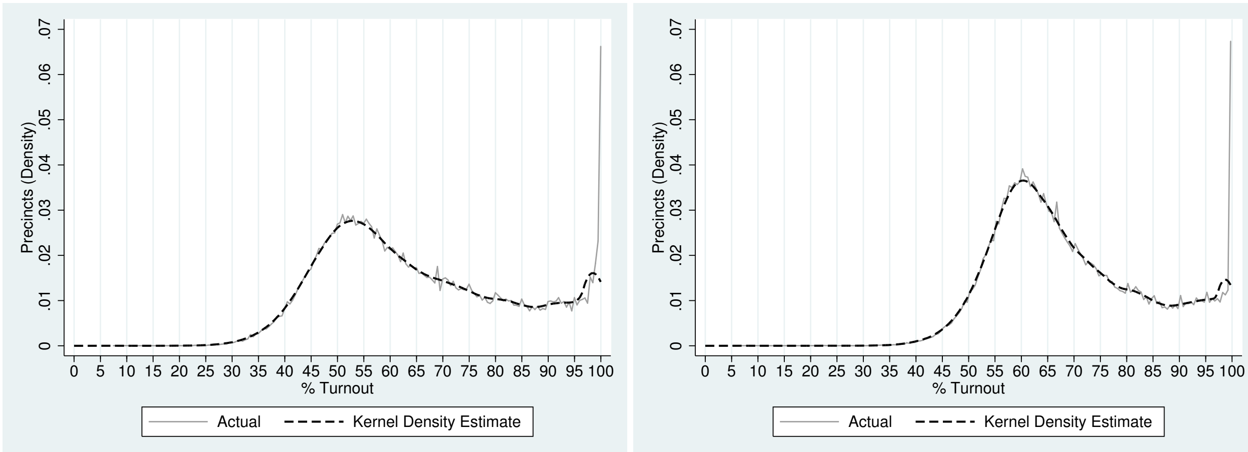
(b) 2012 presidential election

Figure A.28: The distribution of the pooled unit and the first decimal place digits in turnout (after rounding to the nearest multiple of 0.5)

fairly rugged, and like our previous analysis of vote shares, spikes often coincide with multiples of five.

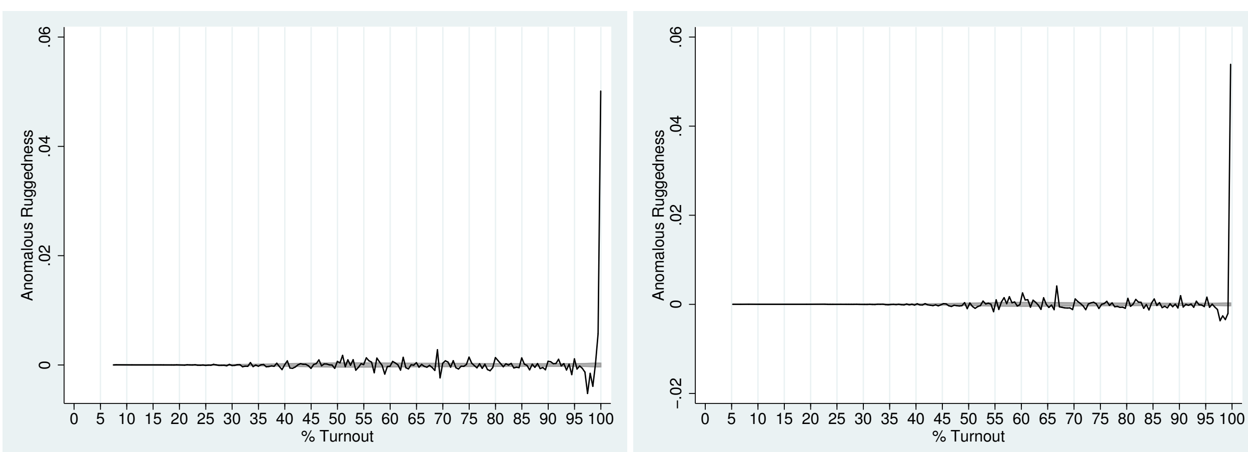
Using our alternative theoretical benchmark, we commute the 95% asymptotic confidence intervals for the kernel density estimate of turnout and treat the empirical observations outside of these confidence intervals as anomalously rugged. Figure A.30 shows significant ruggedness, but the residuals outside of the confidence intervals do not appear to be increasing in turnout.

Figure A.31 presents the absolute value of these residuals, further confirming that ruggedness is not increasing in turnout, except with regards to the large spike around 100%. Since turnout does not become progressively anomalous as it increases, but Putin and United Russia's vote shares do become progressively anomalous as they increase, this suggests that the primary method of fraud used in both elections was stealing from other candidates, presumably second place candidates Zyuganov and the Communist Party.



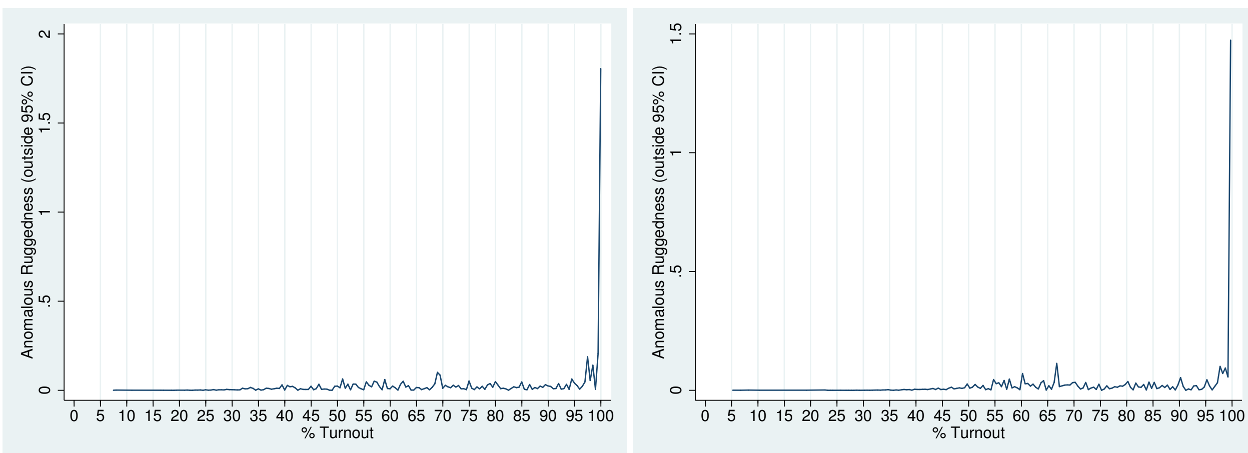
(a) 2011 parliamentary election (b) 2012 presidential election

Figure A.29: The distribution (gray solid line) and kernel density estimate (black dashed line) of electoral turnout



(a) 2011 parliamentary election (b) 2012 presidential election

Figure A.30: The 95% asymptotic confidence intervals (gray area) for the kernel density estimate (black line) of electoral turnout



(a) 2011 parliamentary election

(b) 2012 presidential election

Figure A.31: The difference between the empirical distribution of electoral turnout and the 95% confidence interval of its kernel density estimate

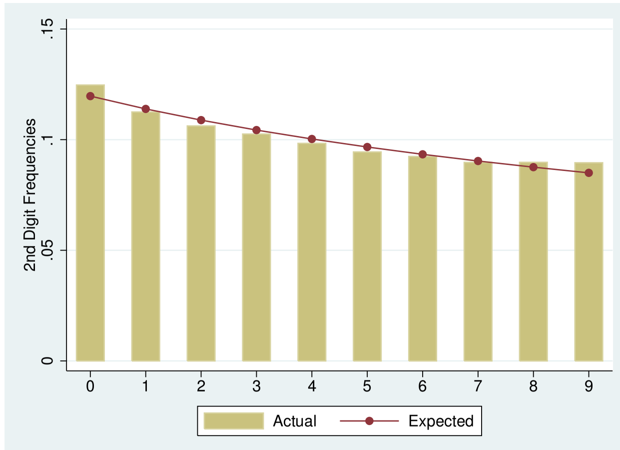
E.1 The 2012 Russian Presidential Election and Benford's Law

According to Benford's Law, the leading digits of many naturally occurring numbers are not uniformly distributed but rather follow a logarithmic pattern, according to which lower numbers occur with greater frequency than higher ones (**??**). While extant research demonstrates the applicability of Benford's Law to population numbers, death rates, and accounting data (**??**), its reliability in detecting electoral fraud is debated (**????**). Here we briefly assess the performance of Benford's Law in detecting fraud during the 2012 Russian presidential election. Since most of the literature on Benford's Law rejects the validity of first-digit models for electoral fraud detection, we conduct a second-digit test (**???**).

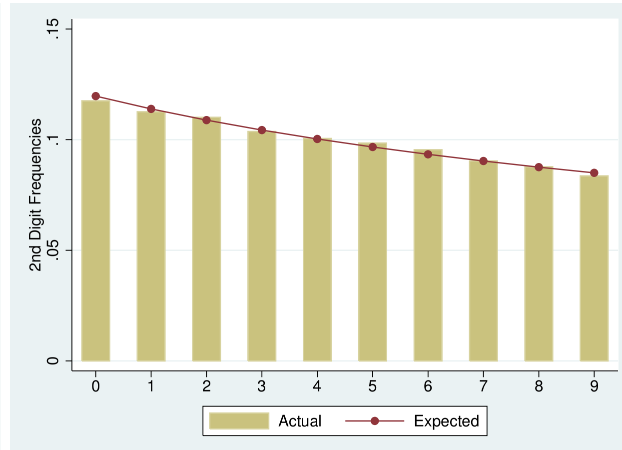
Figure A.32 compares the actual and expected frequency of second digits for each candidate in the 2012 presidential election. In order to test whether deviations of the actual from expected frequency of second digits are statistically significant, we use the Pearson χ^2 test suggested by **?**:

$$\chi^2_{B_2} = \sum_{i=0}^{9} \frac{(d_{2i} - d_2 q_{B_2 i})^2}{d_2 q_{B_2 i}},$$

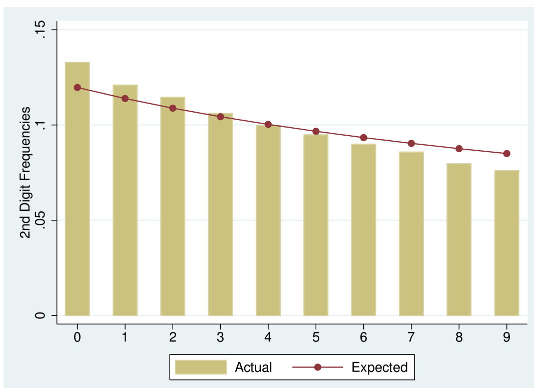
where $q_{B_2 i}$ denotes the expected proportion with which the second digit is i, d_{2i} denotes the actual frequency with which the second digit is i, and $d_2 = \sum_{i=0}^{9} d_{2i}$ denotes the total number of second digits. The test statistic follows the χ^2 distribution with 9 degrees of freedom, and has a critical value of 16.9 for a test at the 0.05 significance level.



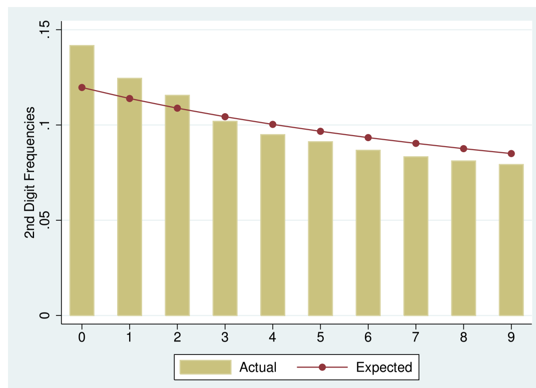
(a) Putin



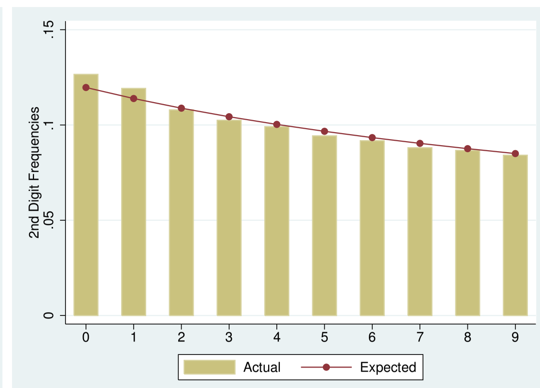
(b) Zyuganov



(c) Prokhorov



(d) Zhirinovsky



(e) Mironov

Figure A.32: The expected and actual frequency of second digits for each candidate's precinct-level results

We find that each candidates' second digits significantly deviate from the distribution implied by Benford's Law, and these deviations are largest for the *worst* performing candidates.[9] Since all accusations of fraud during this election concerned the incumbent Vladimir Putin only, we consider this finding a false-positive. It is most likely due to the fact that the worst performing candidates received close to 0% of the vote in many precincts. As a result, between 19 and 30% of precincts had to be excluded from the analysis of the minor candidates' results, since these candidates' vote counts contained only single digits (i.e. they received fewer than 10 votes). In the remaining precincts, more than 75% of the minor candidates' vote counts were smaller than 100. Thus our case is not suitable for a Benford's law-based test: Benford's law is best applied to data that span several orders of magnitude (**?**) and the $\chi^2_{B_2}$ statistic suffers from false positives when precincts are homogenous and vote percentages are concentrated in narrow range (**?**).

Crucially, tests based on Benford's law do not allow us to evaluate whether the amount of fraud is increasing in the incumbent's precinct-level vote share – a key prediction of our theoretical model that we evaluate in the paper.

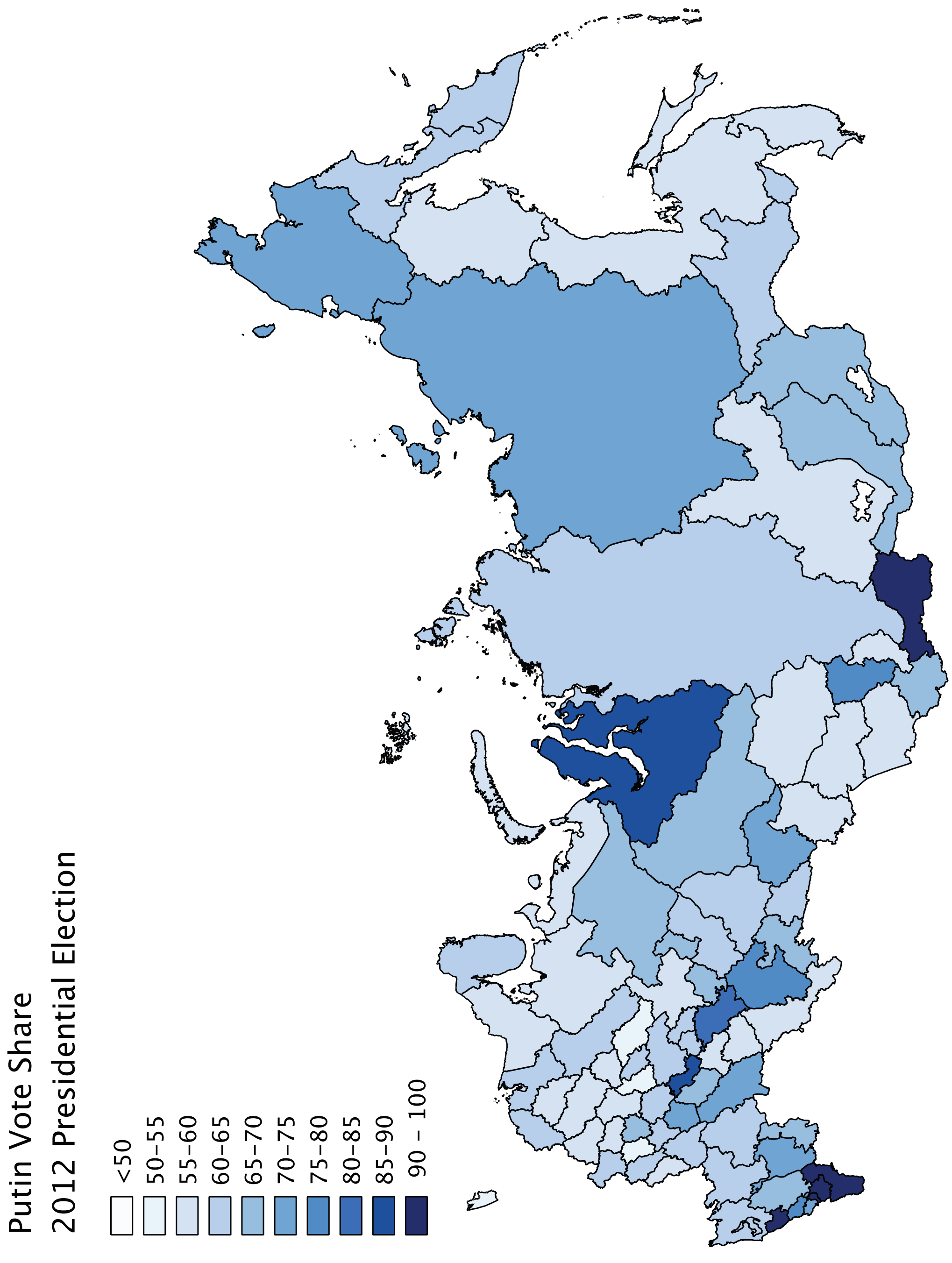
F.1 2011 Russian Legislative Election Vote Shares by Region

[9] All $\chi^2_{B_2}$ statistics exceed the critical value of 16.9: Putin's is 67.8; Zyuganov's is 532.6; Prokhorov's is 9552.6; Mironov's is 11818.1; Zhirinovsky's is 4232.4.

Territory	United Russia	Communist Party	A Just Russia	LDPR
Altai Republic	54.04	21.83	10.46	10.79
Altai Territory	37.89	25.19	16.41	16.89
Amur Region	44.32	19.53	10.46	21.35
Arkhangelsk Region	32.27	20.46	22.37	18.38
Astrakhan Region	61.17	13.48	14.80	8.47
Belgorod Region	52.01	22.79	11.78	9.81
Bryansk Region	50.83	23.64	11.33	10.80
Chechen Republic	99.54	0.09	0.18	0.02
Chelyabinsk Region	51.15	14.88	16.92	11.98
Chukotka Autonomous Okrug	72.56	6.92	5.57	11.60
Chuvash Republic	44.73	21.54	19.36	10.99
City of Moscow	47.45	19.69	12.35	9.62
City of St. Petersburg	35.93	15.57	24.04	10.46
Irkutsk Region	35.43	28.18	13.55	17.58
Ivanovo Region	40.73	22.86	15.84	15.01
Jewish Autonomous Region	49.35	20.31	10.80	16.12
Kabardino-Balkaria	81.94	17.64	0.20	0.08
Kaliningrad Region	37.69	25.96	13.48	14.33
Kaluga Region	41.06	22.26	15.86	14.59
Kamchatka	46.21	17.44	10.27	19.00
Karachay-Cherkess Republic	90.04	8.84	0.47	0.28
Kemerovo Region	65.25	10.67	8.09	12.53
Khabarovsk Territory	38.81	20.85	14.34	20.17
Khanty-Mansiysk Autonomous Okrug	41.75	16.39	14.10	22.94
Kirov Region	35.42	23.02	20.08	16.95
Kostroma Region	31.13	29.21	18.82	16.19
Krasnodar Region	57.06	17.84	10.98	10.62
Krasnoyarsk Territory	37.35	24.02	16.14	17.29
Kurgan region	45.00	19.90	14.67	17.11
Kursk Region	46.41	21.02	14.65	13.67

Territory	United Russia	Communist Party	A Just Russia	LDPR
Leningrad Region	34.38	17.74	25.81	15.13
Lipetsk Region	40.87	23.34	17.06	14.68
Magadan Region	41.61	23.06	11.77	17.60
Moscow Region	575.51	442.84	274.54	249.04
Murmansk Region	32.63	22.17	20.04	18.45
Nenets Autonomous District	36.57	25.17	15.20	17.79
Nizhny Novgorod Region	45.10	29.13	10.73	10.79
Novgorod Region	35.23	19.88	28.58	11.69
Novosibirsk Region	34.35	30.72	12.89	15.94
Omsk Region	40.39	26.07	13.67	14.47
Orel Region	39.72	32.58	11.42	12.47
Orenburg Region	35.38	26.54	17.03	17.14
Perm	37.10	21.49	16.78	18.29
Prenza Region	57.38	20.21	8.82	10.32
Primosrky Krai	33.86	23.87	18.57	19.15
Pskov Region	37.18	25.50	16.65	14.13
Republic of Adygea	61.08	18.94	8.64	7.86
Republic of Bashkortostan	71.22	15.81	5.51	5.26
Republic of Buryatia	49.74	24.70	12.82	9.61
Republic of Dagestan	91.60	7.94	0.19	0.03
Republic of Ingushetia	91.74	2.97	2.34	0.41
Republic of Kalmykia	67.18	18.67	7.30	4.09
Republic of Karelia	32.90	19.64	20.99	18.30
Republic of Khakassia	40.79	24.02	13.90	16.28
Republic of Komi	59.59	13.64	11.62	12.07
Republic of Mari-El	52.93	21.00	10.73	11.87
Republic of Mordovia	92.06	4.57	1.30	1.55
Republic of North Ossetia - Alania	68.75	21.99	6.10	2.26
Republic of Sakha	49.77	16.60	22.09	8.58
Republic of Tatarstan	78.55	10.69	5.35	3.52
Republic of Tyva	86.12	3.97	6.78	2.11
Rostov Region	50.90	21.13	13.44	10.30
Ryazan Region	40.42	23.96	15.32	15.30
Sakhalin Region	42.60	23.81	11.96	16.25
Samara Region	40.16	23.60	14.48	16.04

Territory	United Russia	Communist Party	A Just Russia	LDPR
Saratov Region	65.81	14.00	10.23	7.34
Smolensk Region	36.83	24.64	18.91	15.00
Stavropol Territory	50.05	18.75	12.05	15.60
Sverdlovsk Region	33.55	17.26	25.33	16.42
Tambov Region	67.69	16.72	6.12	7.20
Tomsk Region	38.22	22.81	13.66	18.19
Trans-Baikal Territory	43.95	18.93	14.32	19.48
Tula Region	62.07	15.26	8.56	9.32
Tver Region	38.96	23.55	20.07	11.85
Tyumen Region	62.96	11.88	7.48	14.23
Udmurt Republic	45.80	19.85	11.35	16.85
Ulyanovsk Region	44.19	23.42	15.84	12.77
Vladimir Region	38.87	20.85	21.87	13.14
Volgograd Region	36.00	23.09	22.26	13.48
Vologod Region	34.00	17.08	27.64	15.71
Voronezh Region	50.69	22.13	14.66	8.99
Yamalo-Nenets Autonomous District	72.46	6.63	4.75	13.78
Yaroslavl Region	29.50	24.37	22.98	15.73


Putin Vote Share
2012 Presidential Election
<50
50–55
55–60
60–65
65–70
70–75
75–80
80–85
85–90
90 – 100