Limiting the market for information as a tool of governance: Evidence from Russia

Klaus Ackermann
klaus.ackermann@monash.edu

Department of Econometrics and Business Statistics
Monash Business School
Monash University

MASSIVE | Multi-modal Australian ScienceS Imaging and Visualisation Environment

MONASH University

Motivation

"Shutting down the Internet is a drastic solution that can create problems for the authorities and can hurt the economy. Slowing the Internet connection speed right down is more subtle but also effective as it makes it impossible to send or receive photos or videos. Iran is past master at this. Syria's censors also play with the Internet connection speed, fluctuations being a good indicator of the level of repression in a given region."

Reporters Without Borders 2012

Agenda

Motivation

- Background: Data creation
- Identifying Internet Censorship

Literature

- Media an elections

Data

- World Wide Estimation
- Validation: Tehran, Iran

Estimation

Discussion

Motivation

Human Rights Violation

UN Resolution A/HRC/32/L.20, same rights online as offline

Information is at the core of social choice

Preferences are revealed during elections, Sen, A. (2008)

Updating of institution distorted in Russia

DDoS attacks against blogs which reveal corruption, Enikolopov et al (2013)

The Idea: Identifying Censorship

Hypothesis

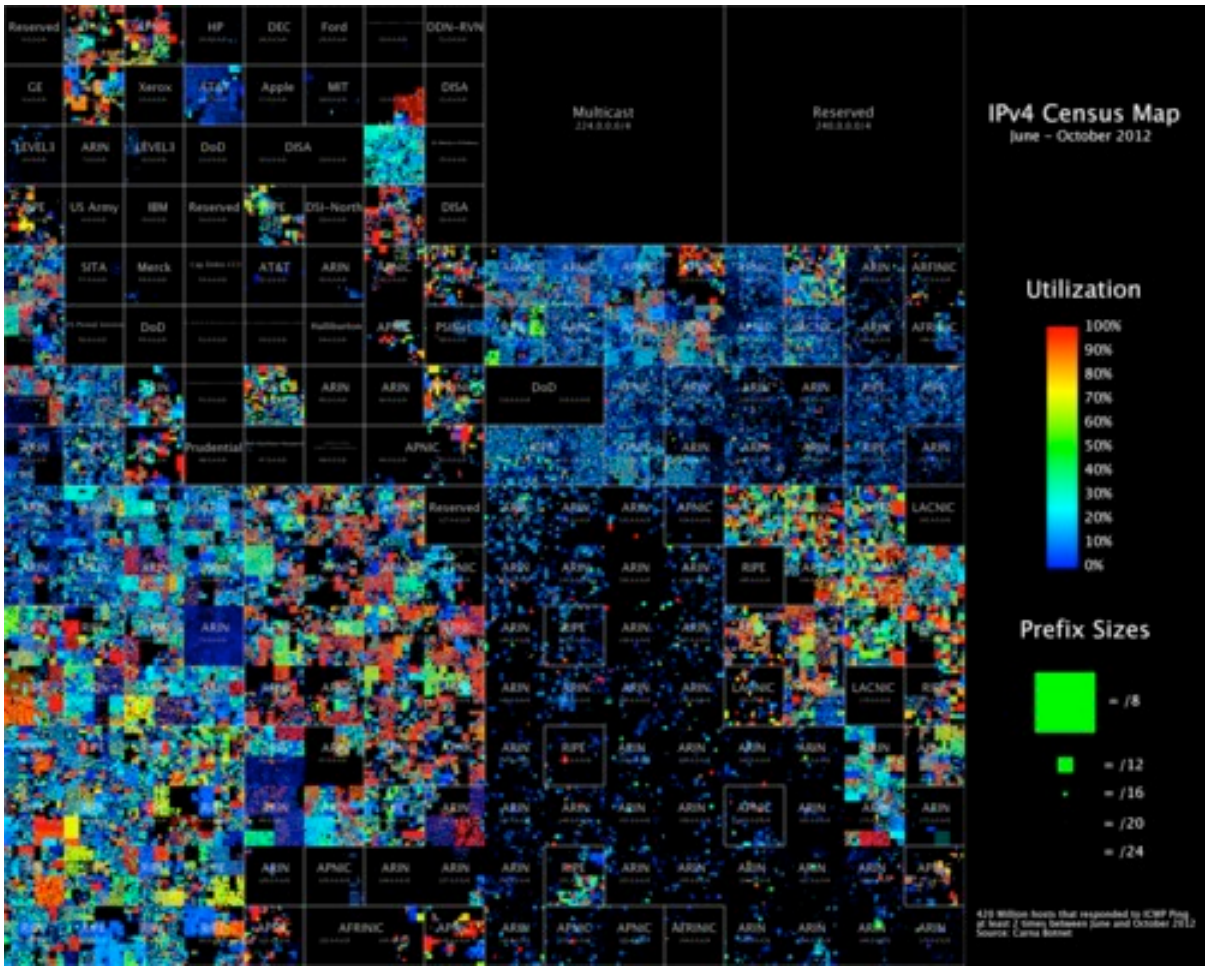
- 1) In autocratic leaning states, the internet is likely not 'free', control is implemented by slowing down internet traffic
- 2) This kind of control will manifest at particular times in which the control of information by the sovereign will favour the interests of the regime (e.g. during elections, political anniversaries, or during times of intense political activism).
- Not looking at a "Twitter-revolution", rather than the transmission of information
- Global data set of Internet Control Message Protocol (ICMP) probes should reveal hidden actions
- Assumption: A sovereign sponsored interference with the Internet in a region should cause a distortion in the response time of ICMP packages. The timing is used for identification.

Side Effect

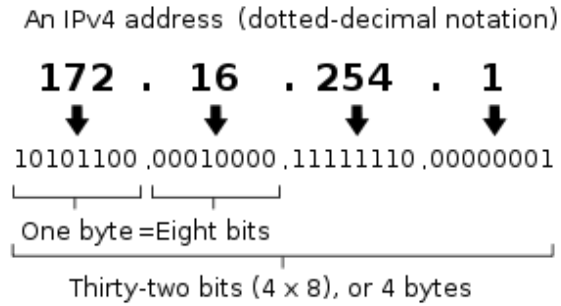
- Slowing down the Internet prevents bypassing of obvious state censorship (VPN, TOR, Screen forwarding etc.)

Can we leverage billions of geo-located ICMP probes to first, identify internet speed tampering in certain regions of Russia, and then, to use these data to estimate the impact of tampering on the 2012 Presidential election results?

Internet Protocol (IP) Addresses, IPv4, and Hilbert Projections



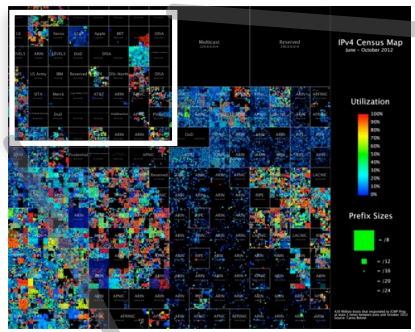
Credit: http://internetcensus2012.bitbucket.org/hilbert.html

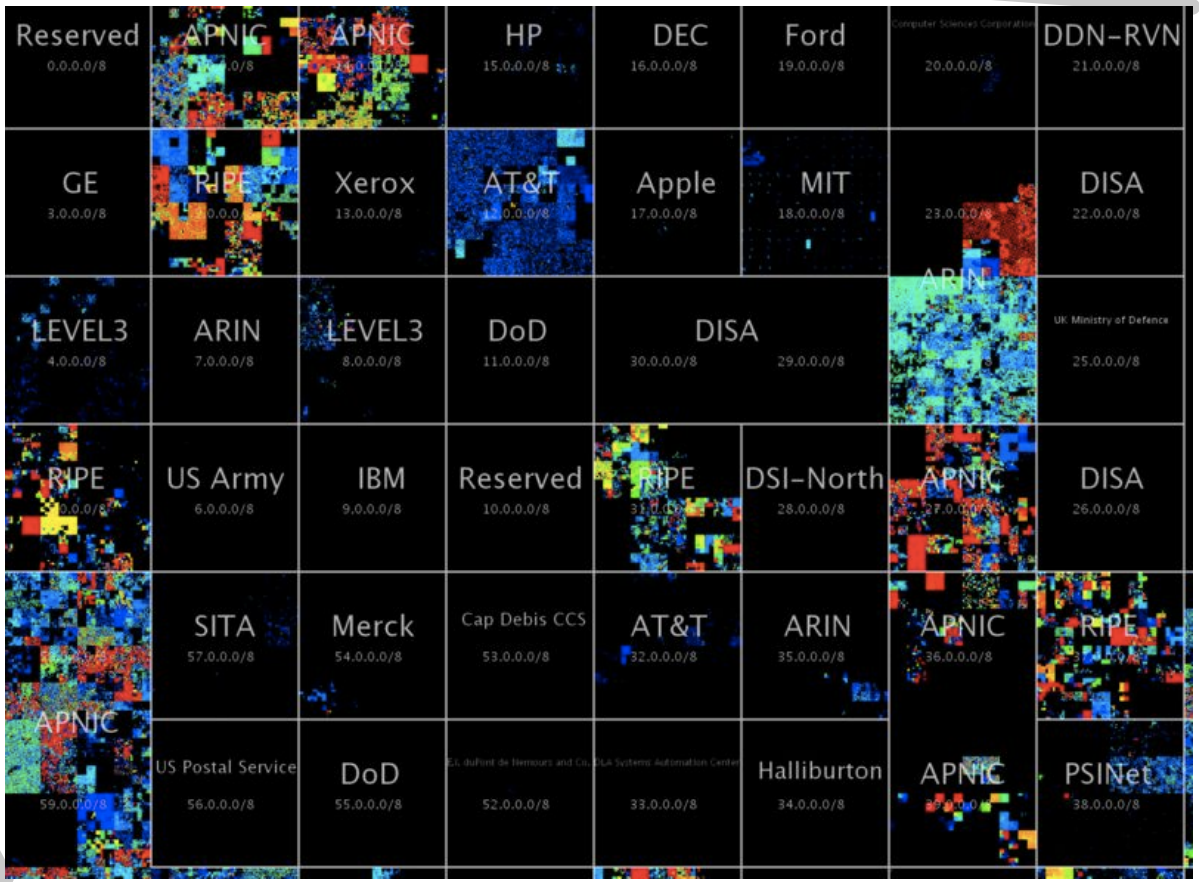


Source: "Indeterminate' (via Wikimedia Commons)

Total possible:
4,294,967,296 (2^{32})
(> 4 billion)

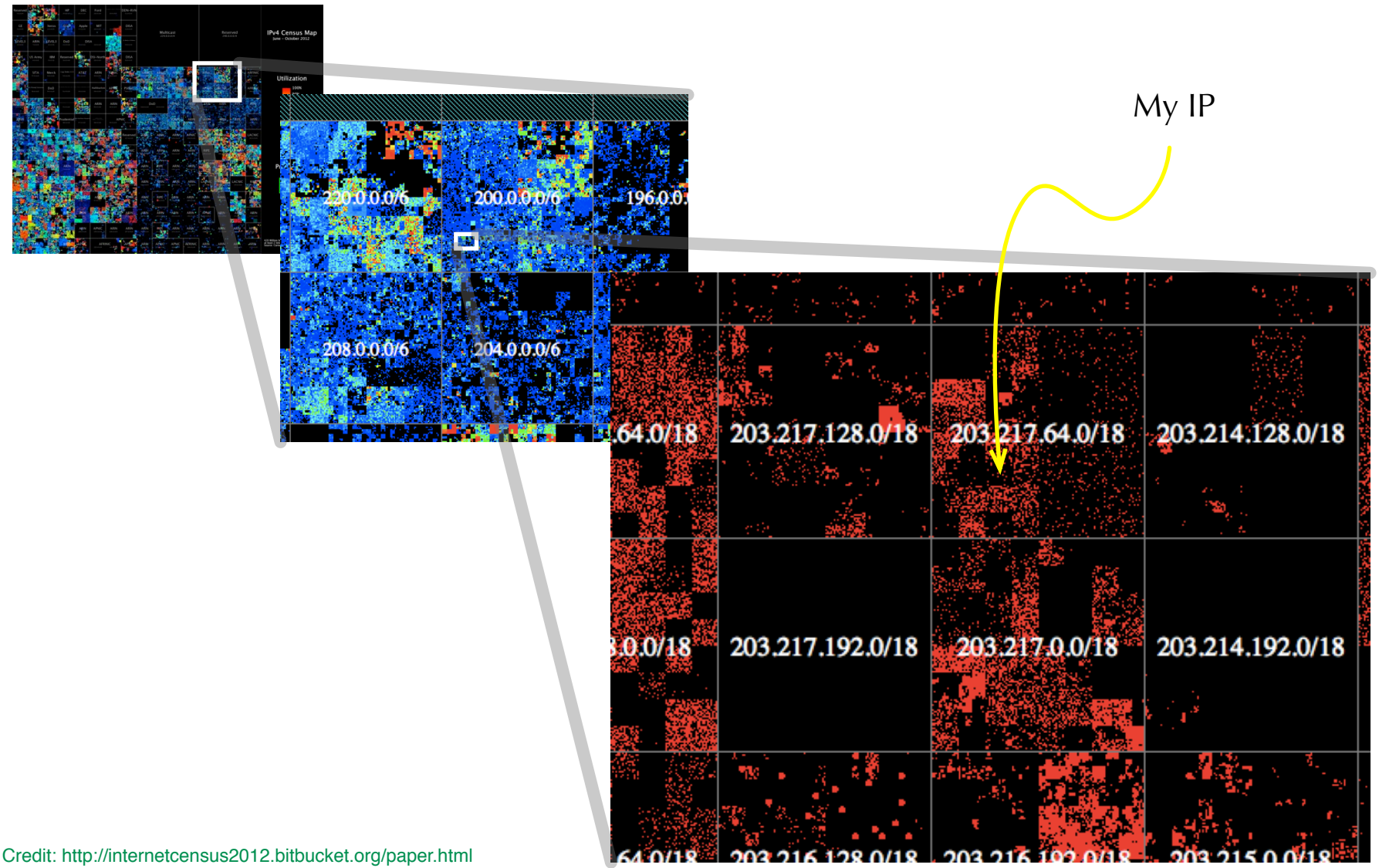
Internet Protocol (IP) Addresses, IPv4, and Hilbert Projections





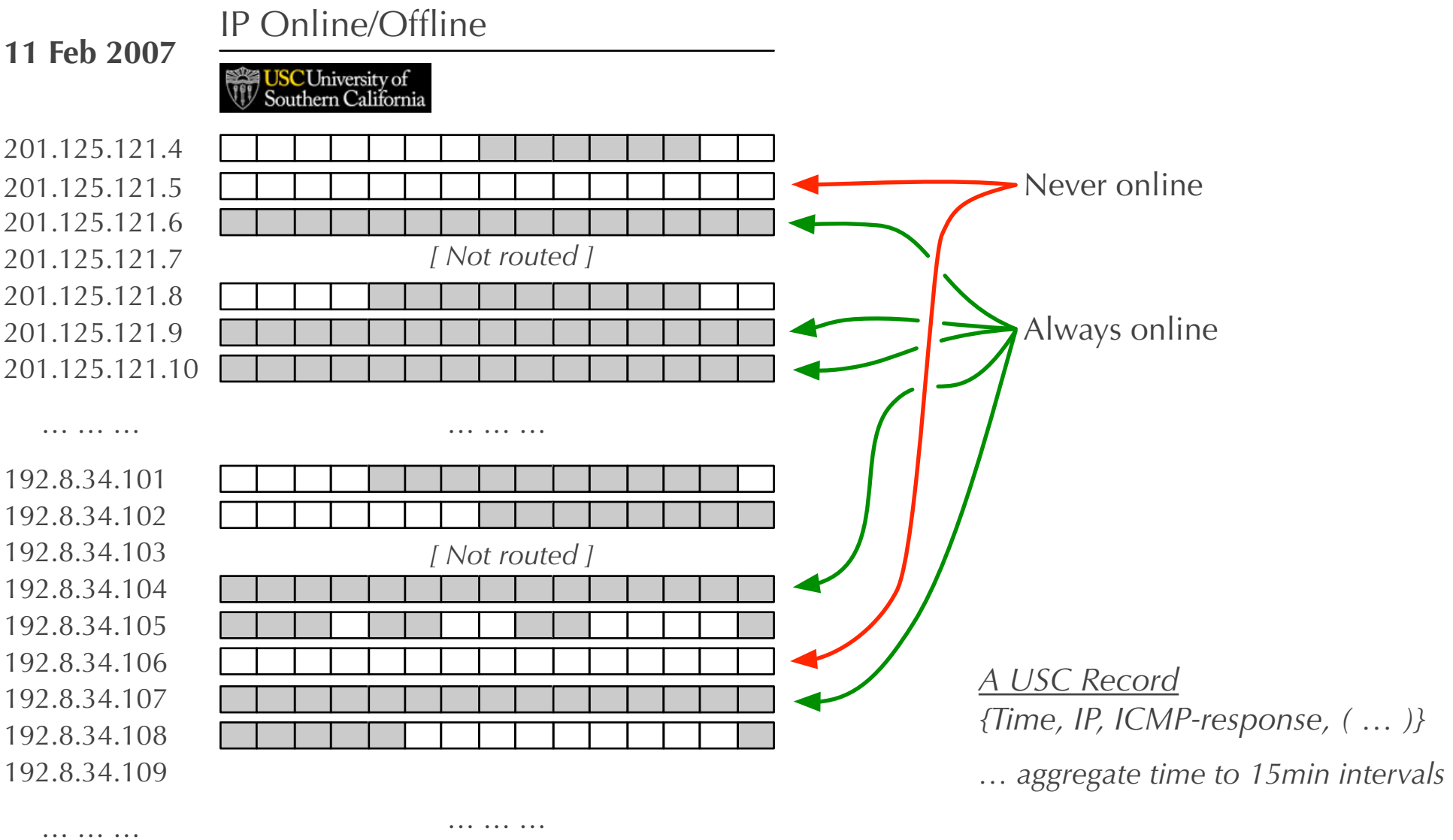
Credit: http://internetcensus2012.bitbucket.org/paper.html

Internet Protocol (IP) Addresses, IPv4, and Hilbert Projections



Credit: http://internetcensus2012.bitbucket.org/paper.html

The Data: USC, Digital Envoy .. to (IP-activity|time|geo-location)



The Data: USC, Digital Envoy .. to (IP-activity|time|geo-location)

12 Feb 2007

IP Online/Offline

USC University of Southern California

201.125.121.4
201.125.121.5
201.125.121.6
201.125.121.7 *[Not routed]*
201.125.121.8
201.125.121.9
201.125.121.10

… … … … … …

192.8.34.101
192.8.34.102
192.8.34.103 *[Not routed]*
192.8.34.104
192.8.34.105
192.8.34.106
192.8.34.107
192.8.34.108
192.8.34.109

… … … … … …

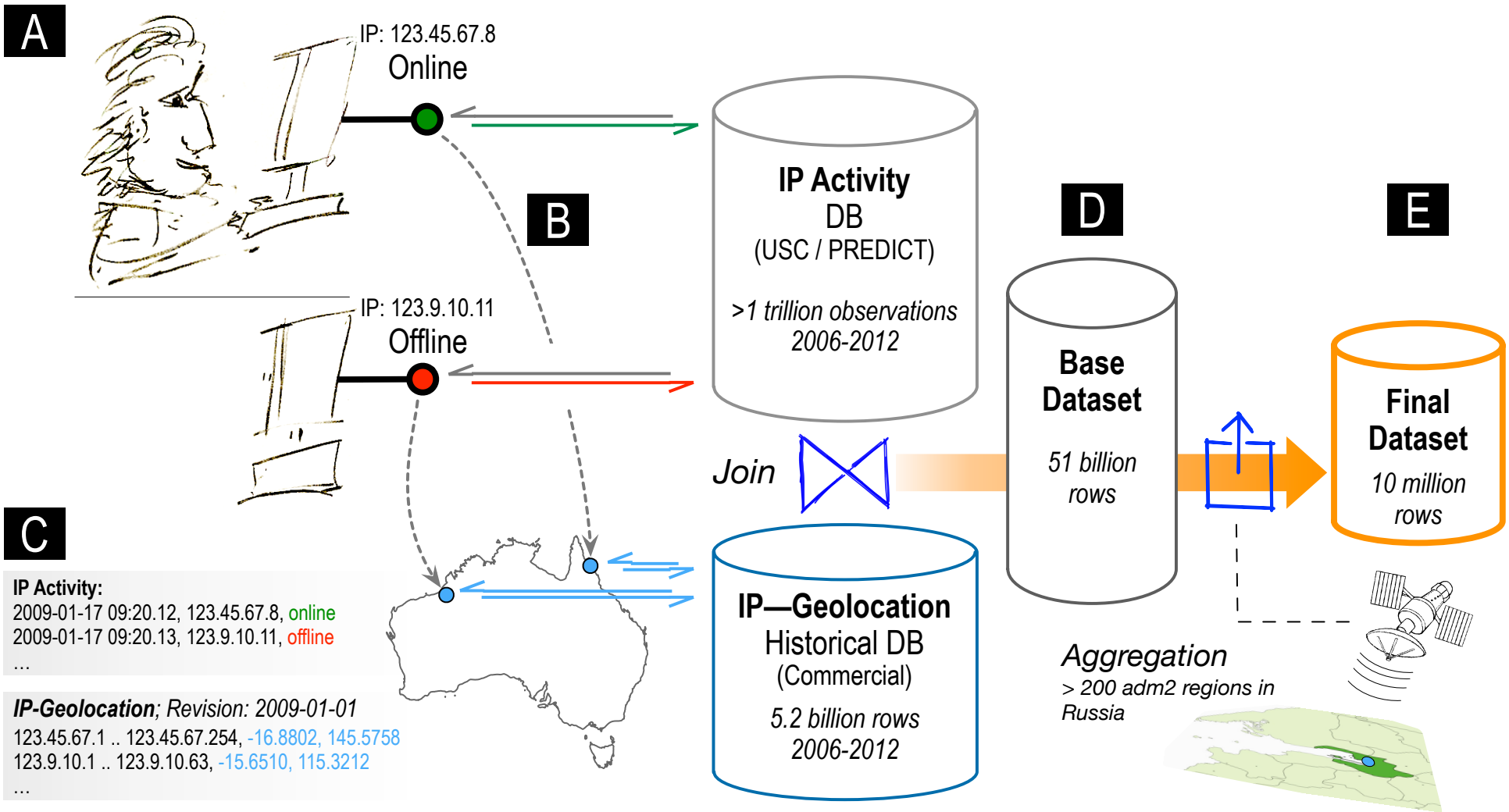
IP —> Location

digital envoy

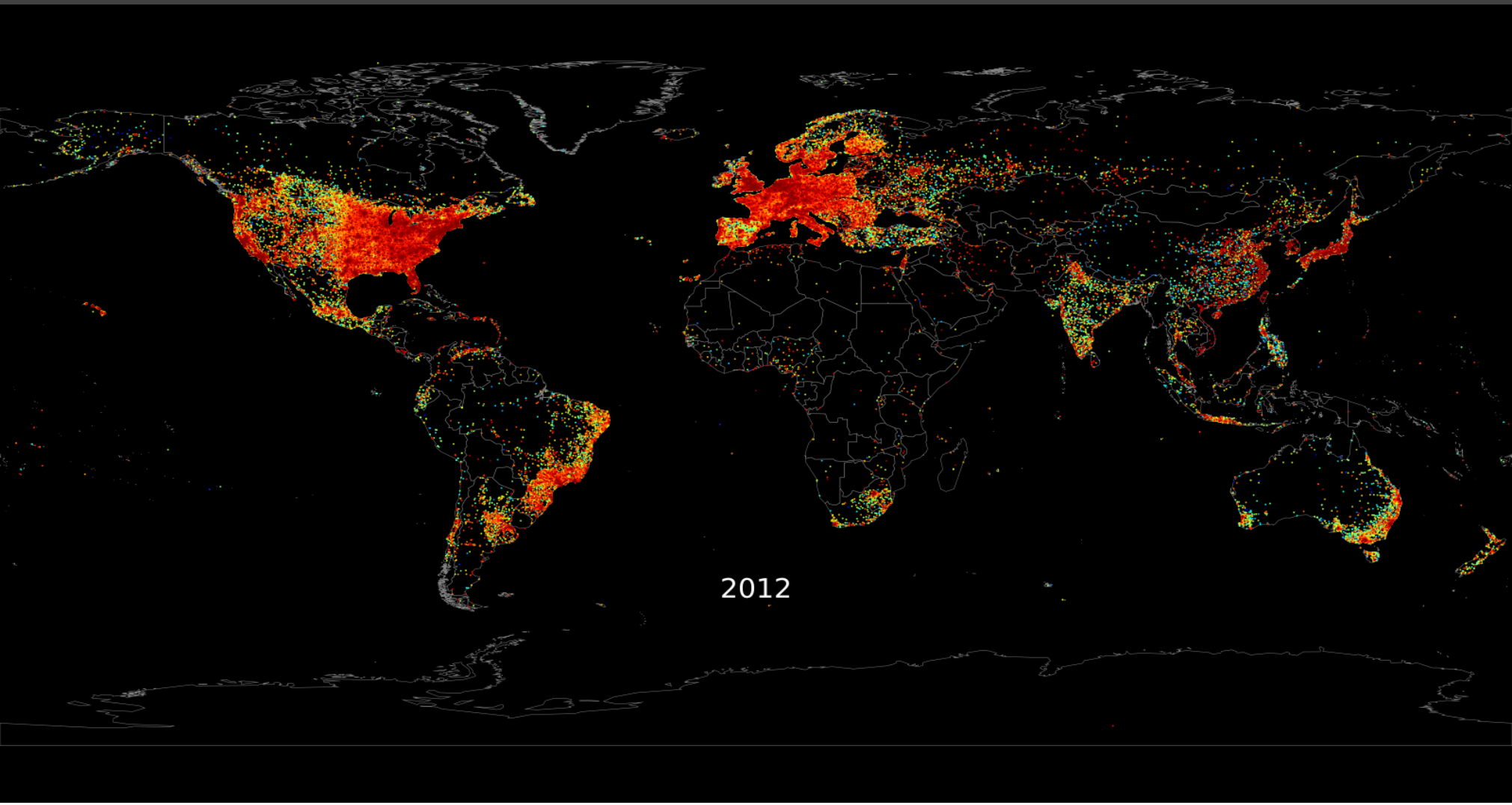
2007.Revision_k

A DE Record
{Time, IP-range, Lat, Lon, (…)}

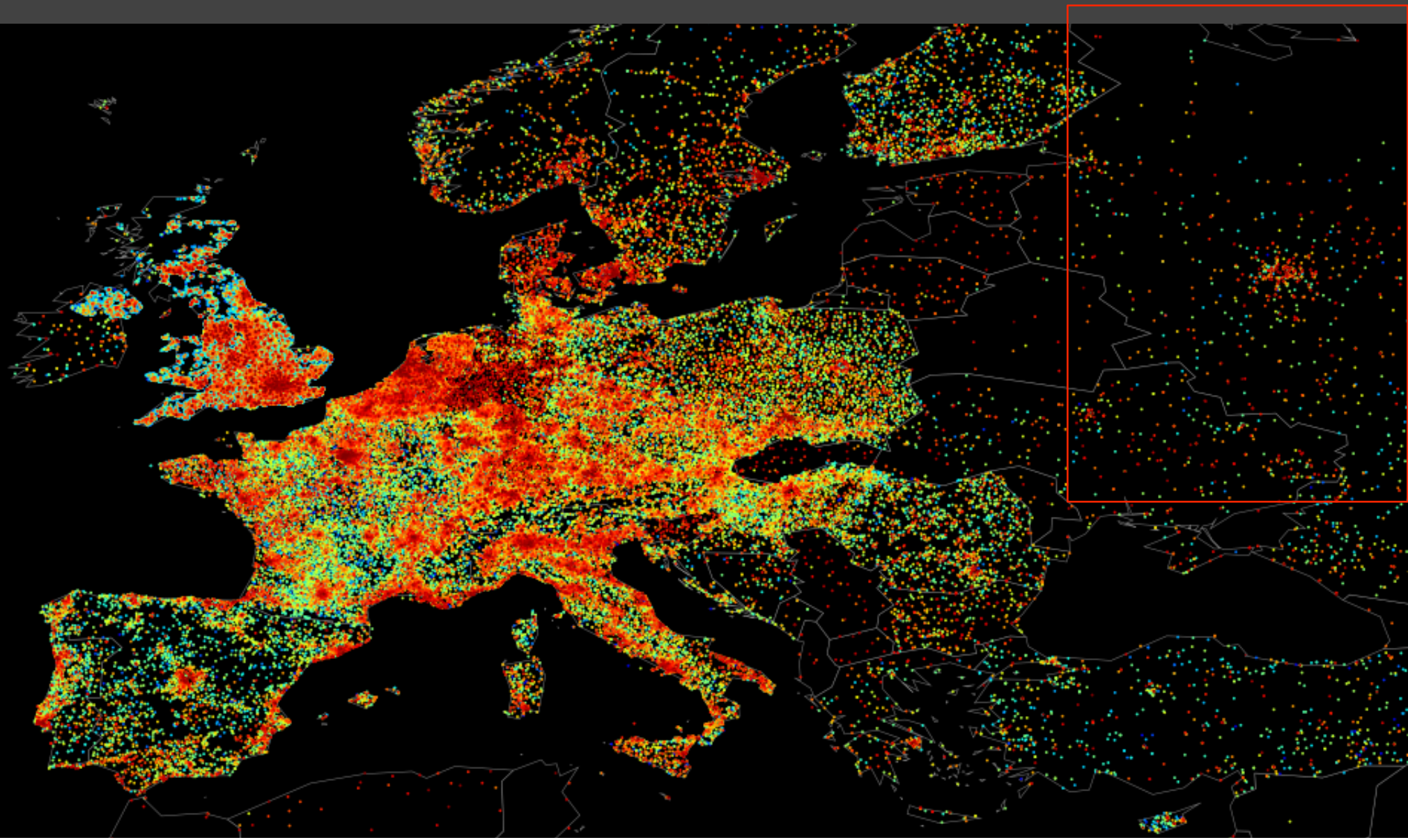
From raw data to revealed global behaviour



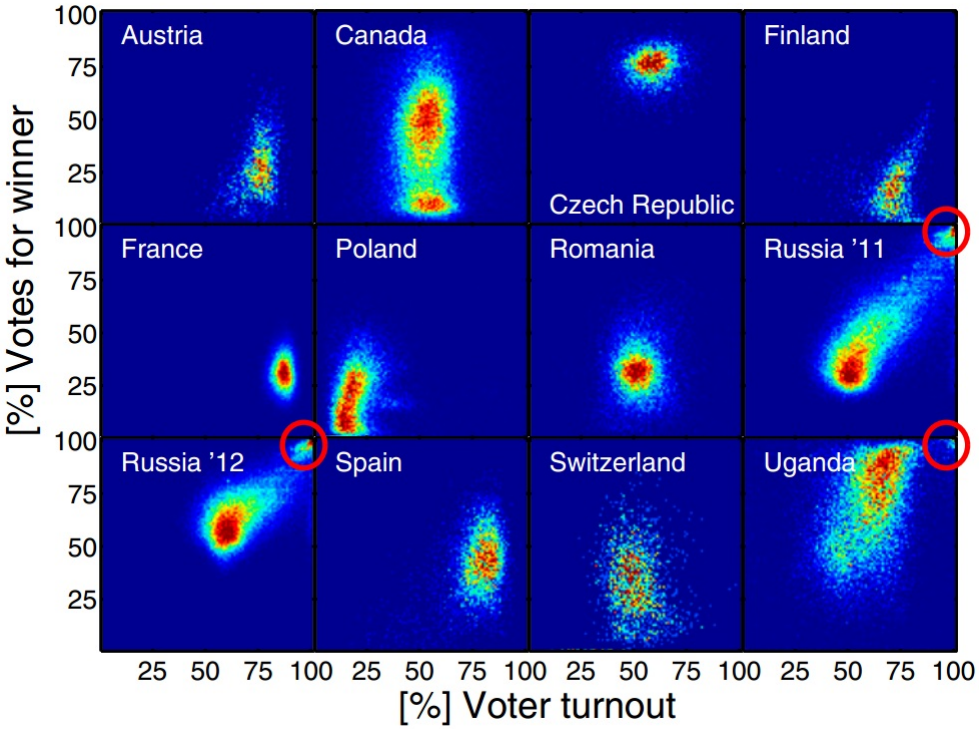
World Wide 2012



Central Europe 2012



Documented Election Fraud in Russia



Klimek, P., Yegorov, Y., Hanel, R., & Thurner, S. (2012). Statistical detection of systematic election irregularities. Proceedings of the National Academy of Sciences, 109(41), 16469–16473.

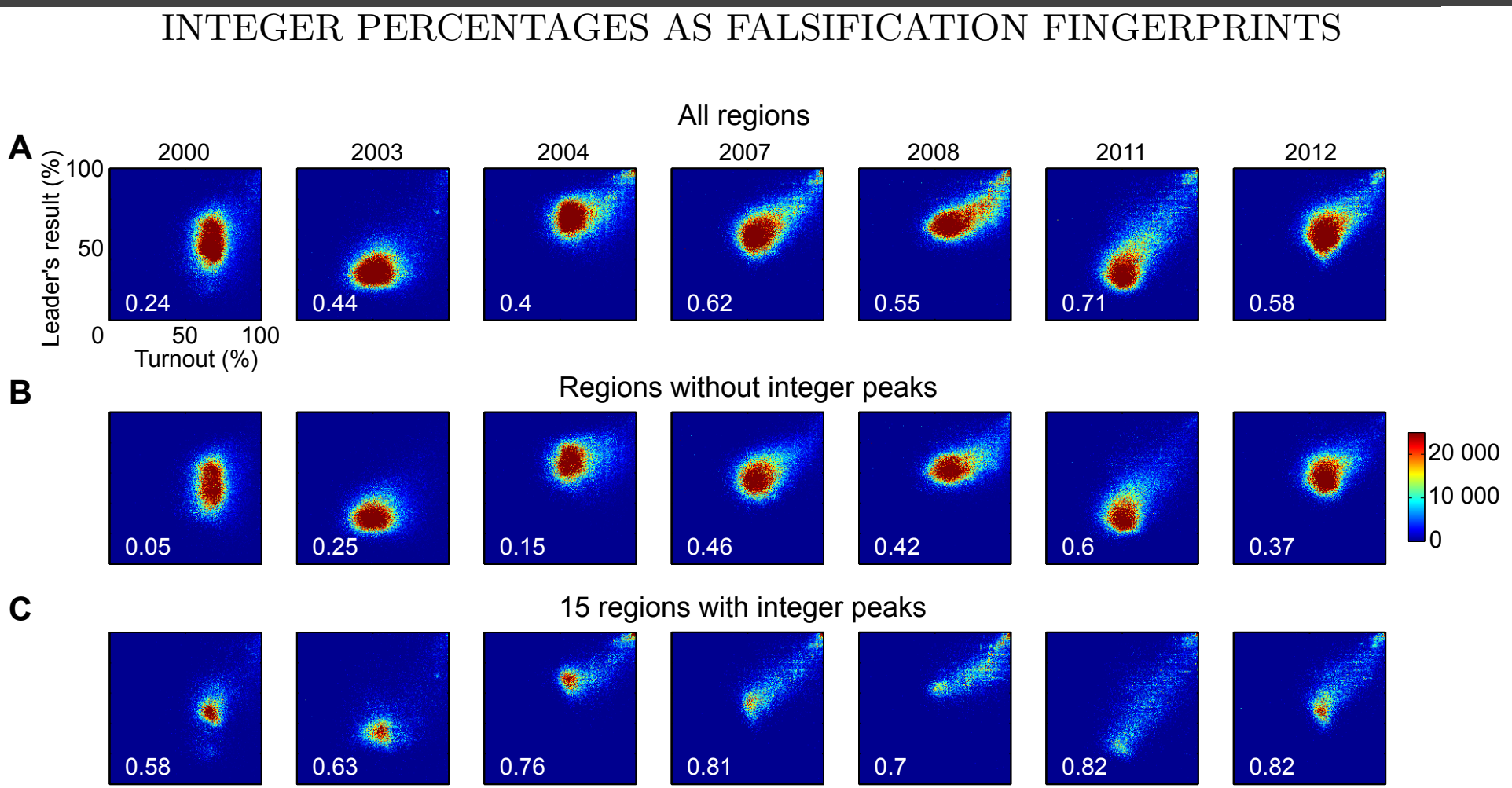
Experiment with Election Observers 2011

Enikolopov, R., Korovkin, V., Petrova, M., Sonin, K., & Zakharov, A. (2012). Field experiment estimate of electoral fraud in Russian parliamentary elections. PNAS, 110(2), 448–452.

Russian Social Media and Protest

Enikolopov, R., Makarin, A., & Petrova, M. (2015). Social Media and Protest Participation: Evidence from Russia. SSRN

Documented Election Fraud in Russia



Kobak D, Shpilkin S. Integer percentages as electoral falsification fingerprints. The Annals of Applied Statistics. 2016;10(1):54-73

Literature on Media

Newspapers

- Increase in print media increases election turn out (Gentzkow et al 2011, AER)
- Competition in the news market does not favour a party (Gentzkow et al 2010, Econometrica)

Television

- One sided media attention increases the likelihood to change the voting preference (Enikolopov et al 2011, AER)
- Fox news effect, broadcasting in favour of the republican party (DellaVigna and Kaplan 2007, QJE)

Internet: A time lag until it affects elections

- Petrova (2008, JPE): Uptake in Internet increases censorship in other news media in autocratic leaning states
- Germany 2004-2008: Increase in Internet penetration reduced the voting turnout (Falck et all 2014, AER)
- Italy 2008-2012: Decreased voter turnout in the short run, long run increase political activism (Durante and Sobbrio 2013, EEA)
- Brazil 2010: Benefit for small parties (Menzes 2015)
- Malaysia 2004-2008: Internet expansion benefited the opposition, closely related study (Miner 2015, JPE)

Detecting Internet Censorship: World Wide Regression

Survey Scan

- Responsive IPs are repeatedly tested every 11 min

Census Scan

- All IP addresses are tested

Observations

- 226,663,017,494 online responses and 1,052,728,499,987 offline (Word Wide)
- 3,201,058,535 online responses and 16,976,018,638 offline (Russia)

$$log(rtt_{icst}) = \alpha_t + \beta_s + \gamma_c + \epsilon_{icst}$$

round trip time rtt, day t, country c, region/isp i,scanning source s

- α_t is a dummy for a day fixed effect
- β_s a dummy for the scanning source
- γ_c a country dummy
- ϵ_{icst} the residual of unexplained variation by region

Survey scan Tehran, Iran: Mobile phone ISP

(1) **14th of February 2012** marked the one year anniversary of protest of the green movement. In the days before the suspected protest against the government in power, international news media, such as the Washington Post, reported limits to the availability of internet bandwidth.

(2) Bid Kaneh explosion.

(3) No censorship noticeable during British Embassy protests.

Average ping times and world traffic corrected difference residuals in Tehran and a mobile phone ISP pre and post the anniversary of the green movement. Based on 6,741,895 survey probes.

Census scan Tehran, Iran: Mobile phone ISP

Average ping times and world traffic corrected difference residuals in Tehran and a mobile phone ISP pre and post the anniversary of the green movement. Based on 596,733 census probes.

Russia Election, Difference in Difference Estimator

Election Data

- Precinct election ballot data matched to sub-region (adm2)
- Duma (parliament) 2011 and Presidential election (2012) (3 months difference)

Samples

- Full sample with regions that had Internet for years
- Robust sample: Excluded regions that previous research identified as fraud

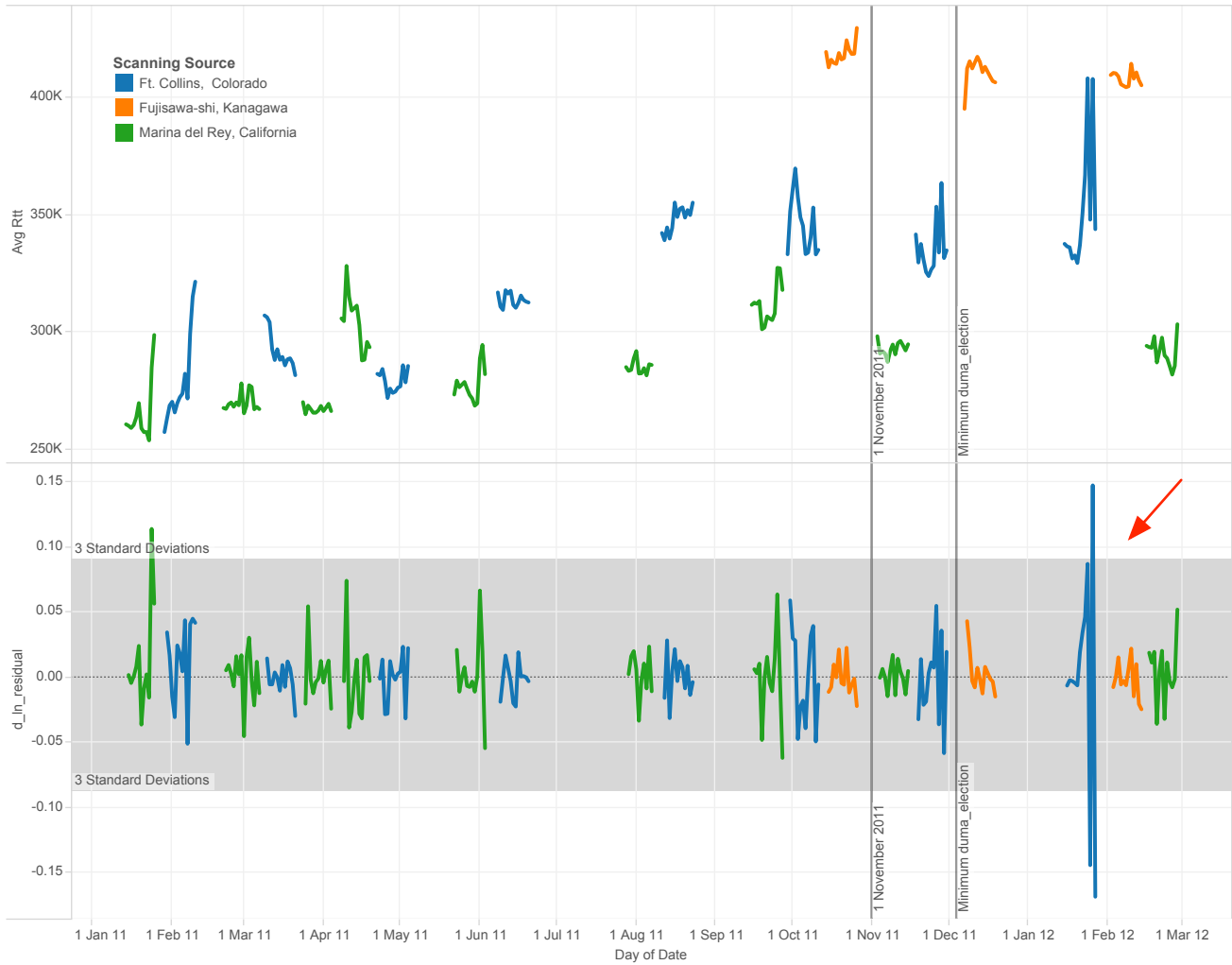
$$Y_{irt}^{j} = \alpha + \beta_1 \cdot censorship_r + \beta_2 \cdot post_t + \beta_3(censorship_r \cdot post_t) + \epsilon_{irt}^{j}$$

$$Y_{irt}^{j} = \alpha + \beta_1 \cdot censorship_r + \beta_2 \cdot post_t + \beta_3(censorship_r \cdot post_t) + \beta_4 \cdot X_{irt} + \epsilon_{irt}^{j}$$

Y_{irt} is the voting share of the party or candidate j, precinct electoral commission i, region r at time t

- $censorship_r$ denotes if a sub-region experiences censorship
- $post_t$ a time trend
- $(censorship_r \cdot post_t)$ is the interaction of interest
- X_{irt} are the electoral controls, precinct election turnout and eligible voters by square kilometre, as well as economic indicators, all in logarithmic scale.

Survey scan Orenburg, Russia



Average ping times and world traffic corrected difference residuals in Orenburg. Based on 4,382,231 survey probes.

Census scan Orenburg, Russia

Average ping times and world traffic corrected difference residuals in Orenburg. Based on 509,908 census probes.

Balance Test: Full Sample

Variable	Control	Treatment	Diff	p-value
Vote Share UR	0.496	0.450	-0.046	0.238
	(0.203)	(0.196)	(0.039)	
Vote Share KPRF	0.200	0.192	-0.008	0.631
	(0.095)	(0.087)	(0.017)	
Vote Share JR	0.132	0.162	0.029	0.056
	(0.073)	(0.084)	(0.015)	
Vote Share LDPR	0.129	0.143	0.013	0.137
	(0.069)	(0.058)	(0.009)	
Election Turnout at PEC	0.619	0.575	-0.045	0.082
	(0.174)	(0.147)	(0.025)	
Election Turnout at PEC (log)	-0.518	-0.583	-0.065	0.109
	(0.278)	(0.237)	(0.041)	
Eligible voters by km2	4.522	4.170	-0.352	0.530
	(2.115)	(2.347)	(0.558)	
IP per capita	-4.431	-4.182	0.249	0.032
	(0.528)	(0.429)	(0.114)	
Active Physicians Rate by 1000	1.510	1.491	-0.019	0.708
	(0.188)	(0.207)	(0.051)	
Primary Income Private Households per head	8.557	8.618	0.061	0.415
	(0.339)	(0.318)	(0.074)	

Note: Standard error of the mean and difference is clustered at sub-region level.

Balance Test: Robust Sample

Variable	Control	Treatment	Diff	p-value
Vote Share UR	0.428	0.391	-0.037	0.206
	(0.138)	(0.123)	(0.029)	
Vote Share KPRF	0.229	0.208	-0.021	0.230
	(0.087)	(0.068)	(0.017)	
Vote Share JR	0.152	0.187	0.035	0.039
	(0.063)	(0.074)	(0.016)	
Vote Shae LDPR	0.148	0.155	0.007	0.411
	(0.060)	(0.049)	(0.009)	
Election Turnout at PEC	0.562	0.544	-0.019	0.335
	(0.136)	(0.114)	(0.019)	
Election Turnout at PEC (log)	-0.603	-0.630	-0.027	0.444
	(0.235)	(0.201)	(0.035)	
Eligible voters by km2	3.954	3.998	0.045	0.945
	(1.962)	(2.081)	(0.646)	
IP per capita	-4.350	-4.010	0.341	0.018
	(0.534)	(0.387)	(0.140)	
Active Physicians Rate by 1000	1.577	1.502	-0.075	0.211
	(0.160)	(0.210)	(0.060)	
Primary Income of Private Households per head	8.528	8.571	0.043	0.603
	(0.333)	(0.259)	(0.082)	

Note: Standard error of the mean and difference is clustered at sub-region level.

Estimation: Robust Sample (free of suspected electoral counting fraud)

	Party share							
VARIABLES	(1) UR	(2) UR	(3) KPRF	(4) KPRF	(5) JR	(6) JR	(7) LDPR	(8) LDPR
Difference in Difference Estimator	0.032*	0.032**	-0.002	-0.002	-0.028**	-0.028**	-0.011*	-0.011*
	(0.018)	(0.016)	(0.009)	(0.009)	(0.014)	(0.013)	(0.006)	(0.006)
Treatment	-0.037	-0.029	-0.021	-0.022	0.035**	0.035**	0.007	0.004
	(0.029)	(0.022)	(0.017)	(0.016)	(0.016)	(0.014)	(0.009)	(0.008)
Timetrend	0.190***	0.181***	-0.029***	-0.019*	-0.114***	-0.106***	-0.070***	-0.071***
	(0.009)	(0.010)	(0.006)	(0.010)	(0.007)	(0.009)	(0.004)	(0.005)
Election Turnout at PEC (log)		0.167***		-0.065***		-0.051***		-0.037***
		(0.028)		(0.018)		(0.011)		(0.012)
Eligible voters by km2		-0.019***		0.009**		0.002		-0.002
		(0.004)		(0.004)		(0.002)		(0.002)
IP per capita		-0.008		-0.000		-0.011		0.012
		(0.019)		(0.013)		(0.011)		(0.009)
Primary Income of Private Households per head		-0.049**		-0.012		0.019		0.001
		(0.021)		(0.016)		(0.013)		(0.013)
Active Physicians Rate by 1000		-0.023		0.012		-0.007		0.016
		(0.042)		(0.026)		(0.027)		(0.012)
Observations	6,007	6,007	6,007	6,007	6,007	6,007	6,007	6,007
R^2	0.442	0.549	0.058	0.146	0.622	0.648	0.395	0.430
Number of sub regions	60	60	60	60	60	60	60	60
F-statistics	216.9	158.1	15.34	9.586	145	80.52	236.9	128.9

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Results based on 2935544223 observations sent from beginning of 2011 to 4th of March 2012

Estimation: Robust Sample Extended with previous elections

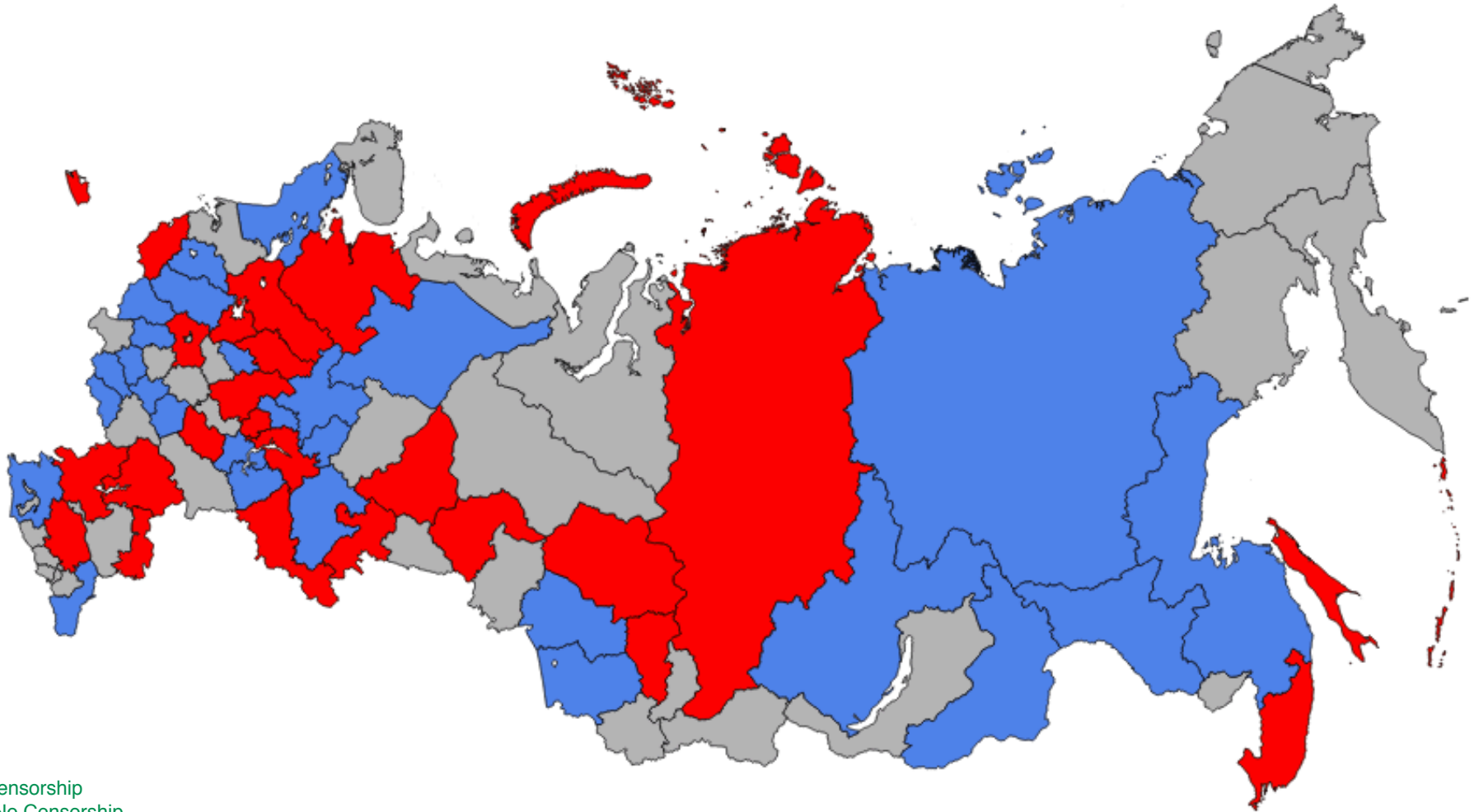
	Party share							
VARIABLES	(1) UR	(2) UR	(3) KPRF	(4) KPRF	(5) JR	(6) JR	(7) LDPR	(8) LDPR
Difference in Difference Estimator	0.040**	0.039**	-0.004	-0.003	-0.036**	-0.035**	-0.010	-0.008
	(0.018)	(0.018)	(0.013)	(0.012)	(0.018)	(0.017)	(0.007)	(0.007)
Election Turnout at PEC (log)		0.120***		-0.062***		-0.029***		-0.026***
		(0.016)		(0.012)		(0.006)		(0.006)
Eligible voters by km2		0.538**		-0.459***		0.009		-0.116
		(0.220)		(0.092)		(0.186)		(0.076)
IP per capita		-0.001		-0.008		0.009		0.010
		(0.021)		(0.010)		(0.014)		(0.007)
Primary Income of Private Households per head		0.064		0.001		-0.008		0.046
		(0.208)		(0.089)		(0.092)		(0.053)
Active Physicians Rate by 1000		-0.168		0.168		0.193		0.012
		(0.174)		(0.102)		(0.153)		(0.073)
Observations	9,993	9,993	9,993	9,993	9,993	9,993	9,993	9,993
R^2	0.688	0.708	0.544	0.563	0.809	0.815	0.492	0.501
Sub region fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sub region time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of sub regions	51	51	51	51	51	51	51	51

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Results based on 2935544223 observations sent from beginning of 2011 to 4th of March 2012

Map of Internet Censored States



RED: Censorship
BLUE: No Censorship
Grey: No consistent internet connection by month since 2006

Alternative Definition: Censorship is assigned to the whole state

	Party share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	UR	UR	KPRF	KPRF	JR	JR	LDPR	LDPR
Difference in Difference Estimator	0.037***	0.015***	-0.010***	-0.001	-0.017***	-0.012***	-0.010***	-0.005**
	(0.006)	(0.005)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Election Turnout at PEC (log)		0.318***		-0.137***		-0.071***		-0.079***
		(0.010)		(0.005)		(0.002)		(0.003)
Eligible voters by km2		-0.006***		0.002***		0.001***		0.001***
		(0.001)		(0.000)		(0.000)		(0.000)
IP per capita		-0.033***		-0.000		0.016***		0.010***
		(0.006)		(0.003)		(0.003)		(0.002)
Primary Income of Private Households per head		0.053*		-0.048**		0.049**		0.000
		(0.032)		(0.020)		(0.023)		(0.011)
Active Physicians Rate by 1000		0.157***		-0.099***		-0.011		0.013
		(0.054)		(0.027)		(0.030)		(0.019)
Observations	286,280	286,280	286,280	286,280	286,280	286,280	286,280	286,280
R^2	0.430	0.593	0.282	0.388	0.581	0.629	0.343	0.415
Region fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of sub regions	1573	1573	1573	1573	1573	1573	1573	1573

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Alternative Definition: State owned Internet Service Provider penetration as continuous treatment

	Party share							
VARIABLES	(1) UR	(2) UR	(3) KPRF	(4) KPRF	(5) JR	(6) JR	(7) LDPR	(8) LDPR
Difference in Difference Estimator	0.052***	0.045***	-0.013*	-0.010	-0.020*	-0.017*	-0.023***	-0.022***
	(0.020)	(0.015)	(0.007)	(0.006)	(0.010)	(0.009)	(0.007)	(0.006)
Election Turnout at PEC (log)		0.318***		-0.137***		-0.071***		-0.079***
		(0.010)		(0.005)		(0.002)		(0.003)
Eligible voters by km2		-0.006***		0.002***		0.001***		0.001***
		(0.001)		(0.000)		(0.000)		(0.000)
IP per capita		-0.034***		-0.000		0.018***		0.011***
		(0.005)		(0.003)		(0.003)		(0.002)
Primary Income of Private Households per head		0.058*		-0.049**		0.048**		-0.002
		(0.032)		(0.020)		(0.023)		(0.011)
Active Physicians Rate by 1000		0.181***		-0.100***		-0.031		0.005
		(0.055)		(0.028)		(0.029)		(0.019)
Observations	286,280	286,280	286,280	286,280	286,280	286,280	286,280	286,280
R^2	0.429	0.593	0.282	0.388	0.579	0.629	0.343	0.416
Region fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region time trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of sub regions	1573	1573	1573	1573	1573	1573	1573	1573

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Thanks to Ruben Enikolopov with the matching to state owned ISPs

Policy Discussion

Compared to Television

- Fox News: 0.4%-0.7% (DellaVigna, S., & Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting, 122(3), 1187–1234. QJE)
- NTV: 1.55% (Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia, 101(7), 3253–3285. AER

Internet is a two way channel

- Censorship impact on the election result: 1.5% - 3.9%
- The goal of internet censorship is not only to prevent the receiving of information but also the spread: probably a lower bound estimate of impact on the election outcome

Other avenues to influence the results

- Modify opinion in other countries as well (e.g. Germany)
- Disallow the opposition (e.g. Yaboloko)
- Integer anomalies (Kobak, D., & Shpilkin, S. (2016). Integer percentages as electoral falsification fingerprints. The Annals of Applied Statistic, 10(1), 54–73).

Outlook

- Rigging of votes on election day is not the only dimension when it comes to electoral fraud. Future research is needed to investigate the effects of internet throttling in other countries.