Supporting Information

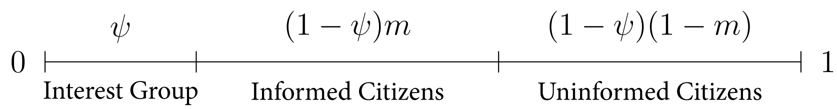
Can You Hear Me Now?:
How Communication Technology Affects Protest and Repression

Following text to be published online.

Contents

A. Model of Coordination, Repression, and Escalation

We model a game between three sets of actors: (1) an interest group considering whether to protest, (2) the government, and (3) a mass of citizens. In a population of measure 1, let ψ belong to the interest group and $1-\psi$ represent other citizens. (We use p to refer to an interest group member and i to refer to a citizen.) Among these $1-\psi$, let $m \in [0,1]$ have access to information about whether a protest happens and any government response. All players know the distribution of the population (ψ and m) and each others' payoff functions.



The sequence of play is as follows:

(1) Before any protest is organized, the government (G) chooses whether to repress in the event of a demonstration ($r \in \{0,1\}$).[1] The government pays a direct cost for deploying repression ($R_G \in \mathbb{R}^1_+$) if a demonstration occurs. This choice is immediately observed by all members of the interest group.

(2) Every interest group member (p) eventually makes two choices: (i) whether to protest, and (ii) what tactic to select.[2] However, before making these decisions, interest group members discuss the plans for a demonstration. Formally, each p receives a vector of S private signals ($\vec{s}_p = \{s_p^1, s_p^2, \ldots, s_p^k, \ldots\}$) about when or where the protest will take place if it occurs. While the distribution of the signals are common knowledge, each p's signals are private and not observed by the government, citizens, or other interest group members. All private signals are independent and identically distributed with each $s_p^k \sim \mathcal{N}(T, 1/\beta_s)$, where T is the actual tactic selected by the protest's organizers. T is an exogenous parameter in this model; it represents

[1] In our one-shot game, allowing the government move first allows it to credibly commit to repressing without complicating the model by introducing repeated play.

[2] We draw upon a recent global game by Little (2015), who presents a tractable approach for modeling protesters' coordination problem. His approach builds upon work by Morris and Shin (2002).

the time or location for the protest chosen by the group's leadership.

Using these signals, each p updates their prior belief $T \sim \mathcal{N}(0, 1/\beta_0)$.[3] Each p's posterior belief about tactics is then

$$E[T|\vec{s}_p] = \mu_p \sim \mathcal{N}\left(\frac{\beta_s \sum_{p=1}^{S} s_p}{\beta_0 + S\beta_s}, \frac{1}{\beta_0 + S\beta_s}\right).$$

To save space, we define $\overline{\beta} = \beta_0 + S\beta_s$ as the precision of this posterior belief.

(3) With this new information in hand, each p then decides whether to protest ($d_p \in \{0, 1\}$) and also selects a tactic ($t \in \mathbb{R}^1$). These choices are observed by the government and *informed* citizens. Furthermore if any p protests and G represses, this repression is then observed by all informed citizens. Any p that protests pays a cost for selecting a tactic that differs from the organizers' plans (T). Furthermore, this cost is larger when the government has chosen to repress demonstrators: botched coordination is especially costly when p shows up at the wrong time and faces the police without compatriots. Specifically, we assume the cost function $(k+rR_p)(t-T)^2$, where $k \in \mathbb{R}^1_+$ scales the cost of botched coordination even absent repression, $r \in \{0, 1\}$ is the government's choice of repression, and $R_p \sim U[0, 1]$ is each p's cost to being repressed. Should they succeed, each p benefits from the policy concession, receiving $c \in \mathbb{R}^1_+$. Let $\mathcal{D}$ be an indicator for whether at least one interest group member chooses to protest ($\mathcal{D} = \mathbb{1}(d_p = 1 \text{ for some } p)$). If $\mathcal{D} = 0$, then the game ends with the government retaining the concession without incurring the cost of repression, all p getting nothing, and all citizens receiving their reservation value $q \in \mathbb{R}^1_+$.

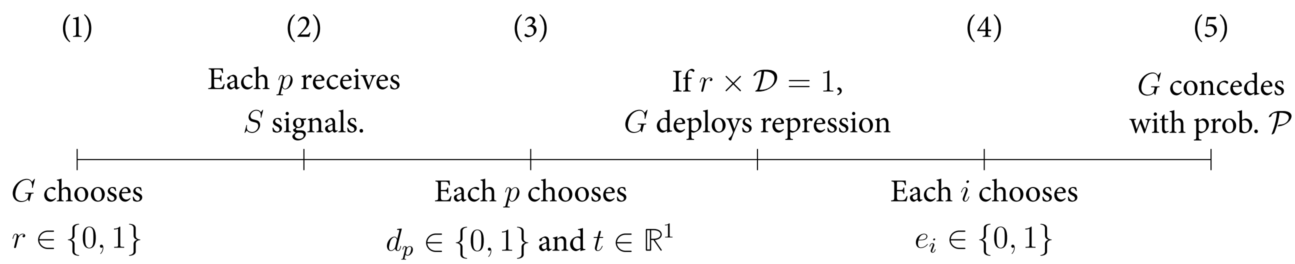
(4) If a protest does occur ($\mathcal{D} = 1$), each citizen decides whether to punish the government ($e_i \in \{0, 1\}$). All uninformed citizens remain ignorant and receive zero utility from indiscriminately punishing the government. Each informed citizen responds differently upon witnessing repression — some may be outraged, others cowed. If i is informed and the government deploys repression ($r \times \mathcal{D} = 1$), they receive $v_i \in \mathbb{R}^1$ (which is distributed according to the cumulative

[3] The precision parameters β_0 and β_s are assumed to be known to all players.

distribution function $F\{\cdot\}$) for choosing $e_i = 1$ and their reservation value (q) for $e_i = 0$.[4]

(5) The game ends with a lottery in which the government concedes with a probability that increases in the measure of protesters and citizens that punish ($\mathcal{P} = \text{measure}[p \mid d_p = 1 \cup i \mid e_i = 1]$). For convenience, we assume that the probability of concession is simply equal to this measure $\mathcal{P}$. If the government prevails, it keeps the concession c. However, if the protest succeeds, then the concession is granted to the interest group members. Regardless of the protest's success, informed citizens get $(r \times \mathcal{D})v_i$ if they punish and q if not.

The following figure summarizes the timing of the game:



We can now state each player's complete strategy: $G : r \to \{0, 1\}$; $p : \{r, R_p, \vec{s}_p\} \to \{0, 1\} \times t$; and informed $i : \{v_i, r \times \mathcal{D}\} \to \{0, 1\}$. We also define each player's expected payoffs both in words and using the notation introduced above:

- G: $E[u_G(r)] = E(\text{Concession}) - \text{Cost of Deploying Repression} = c(1 - \mathcal{P}) - r\mathcal{D}\ R_G$
- p: $E[u_p(d, t)] = \mathbb{1}(\text{Protest}) * E(\text{Concession} - \text{Coord. Cost}) = d_p[c\mathcal{P} - (k + rR_p)E(t - T)^2]$
- Informed i: $E[u_i(e)] = \mathbb{1}(\text{Punish}) * \text{Outrage} + \mathbb{1}(\sim \text{Punish}) * \text{Res. Value} = e\ r\mathcal{D}\ v_i + (1 - e)q$.
- Uninformed i: $E[u_i(e)] = \mathbb{1}(\sim \text{Punish}) * \text{Res. Value} = (1 - e)q$.

A.1 Equilibrium Characterization and Comparative Statics

We derive the equilibrium through backwards induction, starting with the citizens' decision to escalate, then the interest group members' decision to protest, and, finally, the government's initial choice of repression.

Citizens react to what they see transpire in the streets. Did the government repress demon-

[4]This assumes that i does not directly value the concession. We can relax this assumption and allow c to enter i's utility, increasing the measure of citizens that escalate for any level of repression.

strators, and is the citizen angered (scared) enough by this repression to want to take (avoid) action? The case studies and survey evidence cited above suggest that witnessing repressive acts can mobilize some citizens to sympathize with protesters. A citizen chooses to punish the government if their outrage, upon observing repression, exceeds their payoff from remaining neutral. If no repression occurs (or if citizens are uniformed), then nothing incites citizens, and no escalation occurs.[5]

Second, interest group members have to evaluate whether the expected value of the policy concession exceeds the costs of protesting. Their expected benefits ($\mathcal{V}$) from protesting depend on what proportion of group members protest ($\psi\overline{R}$) and what proportion of citizens (if any) choose to punish ($\mathcal{E}$). In short, the more people that demonstrate or punish, the better the chances that the government concedes.[6] Each potential protester's cost to demonstrating depends on their choice of tactic. As this choice is symmetric and does not depend on others' actions, we can immediately solve for each interest group member's optimal tactic: they simply choose their best guess about t based on the signals they received, i.e., their posterior belief (proof in Appendix B.1). This yields the following expected utility to protesting for every protester: $E[u_p(\mu_p)] = d_p[\mathcal{V} - (k + rR_p)/\overline{\beta}]$.[7] As is already apparent from this expression, the costs of coordination decrease as p receives more information about the logistics of protest (because $\overline{\beta}$ is increasing in the number of signals, S).

Finally, the government has to decide whether to repress. The government wants to repress only when the expected deterrent or demobilizing effects of repression outweigh the costs associated with alienating citizens. We define $\mathcal{E}$ as the increased probability that the government is forced to

[5]Alternatively, one could allow citizens to experience outrage even absent repression. This amendment would allow m to affect protest even absent repression.

[6]We define the expected value of the concession as $\mathcal{V}$, which is equal to $c\psi$ if no repression occurs and $c[\mathcal{E} + \psi\overline{R}]$ if the government intervenes, where $\mathcal{E}$ (defined below) represents the measure of citizens that punish after observing repression, and $\overline{R}$ identifies the interest group member that is indifferent between protesting and not. $\overline{R} = \arg_{R_p} \left\{ c[\mathcal{E} + \psi R_p] = (k + rR_p)(t - T)^2 \right\}$.

[7]The expectation simplifies because $E(\mu_p - T)^2 = 1/\overline{\beta}$. Conveniently, $E[(\mu_p - T)^2]$ is simply the variance of the posterior μ_p or $1/\overline{\beta}$.

concede if escalation occurs.[8]

The preceding paragraphs are summarized in the following proposition:

Proposition 1. ***(Equilibrium Characterization)*** *A unique Perfect Bayesian Equilibrium exists. In it, the following properties hold:*

(i) Protests never occur if the expected value of the concession, absent any escalation by other citizens, does not exceed the cost of coordination ($\mathcal{V} < k/\overline{\beta}$).

(ii) However, if this first condition does not hold, the government faces the possibility of protest and represses if the deterrent value of repression exceeds the direct cost of repression, as well as the cost of any escalation ($\psi(1-\overline{R}) \geq \mathcal{E} + R_G/c$).

(iii) An interest group member protests if the expected value of the concession exceeds their costs of coordination and repression. If this is not true for any member of the interest group, then no protest occurs. (An interest group member p whose cost to being repressed is R_p protests if $\mathcal{V} \geq (k + R_p)/\overline{\beta}$.)

(iv) An informed citizen punishes the government if he observes repression and his outrage exceeds his reservation payoff ($v_i \geq q$). Uninformed citizens never punish.

Proof: See Appendix B.2. □

We focus on two comparative statics. The equilibrium changes if we allow interest group members to more intensely communicate. If each member of the interest group receives more signals (increasing S), the possibility of mis-coordinating diminishes. When an interest group member is more confident that he or she will choose the correct tactic, their costs to protesting decline regardless of the government's choice of repression. This makes protest more likely.

Second, what if we expand the audience that observes the government's choice of repression (expand m)? Increasing the proportion of informed citizens amplifies the government's downside risk if it represses, making it less likely to intervene. As the expected level of repression falls, so too

---

[8]Let $\mathcal{E} = (1-\psi)m(1-F\{q\})$. This is simply the measure of informed citizens, whose outrage exceeds their reservation value (i.e., for whom $v_i > q$).

does the cost of protesting for interest group members.

These results are now collected in the following proposition:

Proposition 2. ***(Comparative Statics)*** *The unique PBE, characterized in Proposition 1 above, has the following comparative statics:*

(i) Direct Effect: As the intensity of communication among intrest group members increases, these individuals become more confident that they will coordinate on the correct tactic, lowering their expected costs to protesting. More technically, an interest group member's posterior belief concentrates around the truth as they receive more signals.

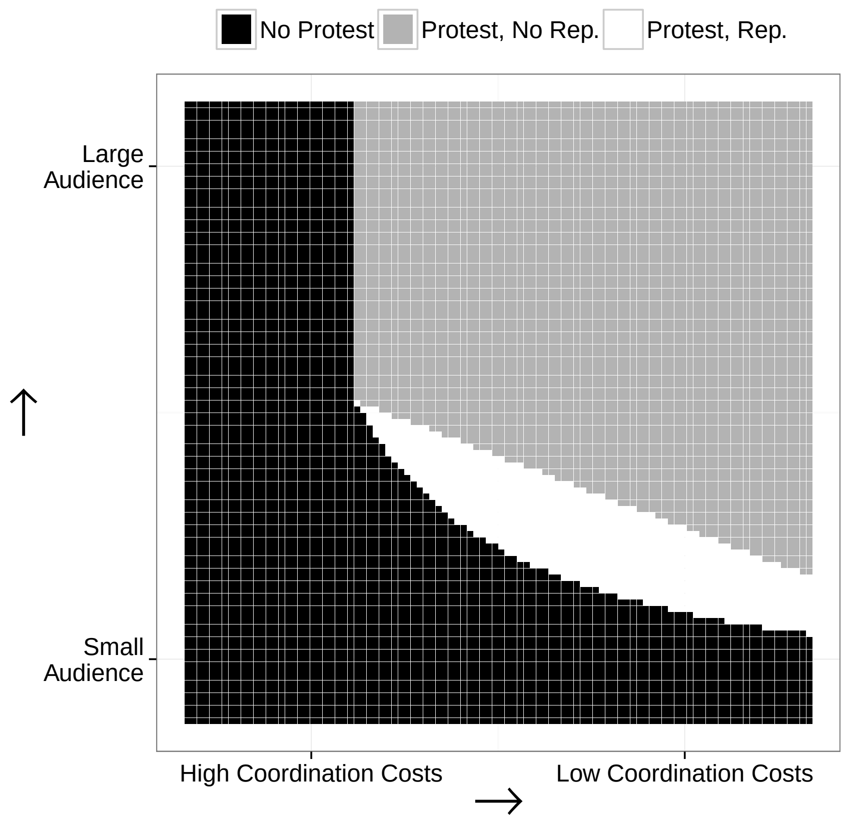
(ii) Indirect Effect: Repression is less likely as the audience of informed citizens (m) increases (so long as $\mathcal{V} \geq k/\overline{\beta}$). As the likelihood of repression falls, so too do the costs of protesting.

Proof: See Appendix B.3. □

A simple way to present these comparative statics is to map out the equilibrium reached for different costs to coordinating (which are a function of $\overline{\beta}$) and audience sizes (m), holding the other parameters fixed. As is apparent in figure A.1, if coordination costs are too high protest is not possible. However, below this threshold, the likelihood of protest is increasing as coordination costs fall and the audience size increases.[9]

[9] While our theory delivers comparative statics about the incidence of protest, it can also be potentially useful to identify protest size, assuming that protests that are joined by outraged informed citizens are larger in size. Figure A.1 presents two regions with equilibrium protest, each corresponding to small or large protest. In the white region, protest occurs, the government represses, and outrage citizens join in; these are large demonstrations. In the light gray region, protest also occurs, but the government is deterred from repressing by potential citizen backlash; these protests remain small. We don't investigate protest size in the empirical analysis, but note that the model has the potential to shed light on this feature of protest.

Figure A.1: Equilibrium as Coordination Costs, Audience Size Change
Lowering coordination costs and increasing audience size increases Pr(Protest).



We map the equilibrium reached at different values of $\overline{\beta}$ and m, the two parameters in our model that we relate to cell phone access. To create this figure, we set ψ = .05, c = 1.5, $F\{q\}$ = .9, R_G = 0.001, and k = .2.

B. Proofs

B.1 Proof of Protester's Tactical Decision

Proof. Each protester chooses the tactic that maximizes her expected utility, given her signals, $\vec{s}$:

$$
\begin{aligned}
t^* &= \arg\max_t E_T[u_p(t)|\vec{s}] \\
&= \arg\max_t E_T[c\mathcal{P} - (k + R_p)(t - T)^2|\vec{s}] \\
&= E_T[T|\vec{s}] = \mu_p
\end{aligned}
$$

□

B.2 Proof of Proposition 1 (Equilibrium Characterization)

Proof. An informed citizen i never wants to punish if no repression occurs, as $0 < q$. However, if i observes repression and $v_i > q$, then they punish the government. This implies that the government alienates a proportion $(1-\psi)m[1-F\{q\}]$ of citizens by choosing to repress.

If the government does not repress at all, no citizens punish, and an interest group member p only protests if $c\psi \geq k/\overline{\beta}$.

Suppose that $c\psi < k/\overline{\beta}$. The government can ensure their maximum payoff c by not repressing. If the government represses, then $(1-\psi)m[1-F\{q\}]$ of citizens punish, reducing the government's expected payoff to $c(1-\psi)m[1-F\{q\}] \leq c$. Thus, if $c\psi < k/\overline{\beta}$, then no protests occur, the regime never represses, and no citizens punish.

Suppose instead that $c\psi \geq k/\overline{\beta}$. If the government represses, then it is punished by $(1-\psi)m[1-F\{q\}]$ citizens, and p wants to protest if their expected utility to protesting is greater than their status quo payoff:

$$\underbrace{c[(1-\psi)m[1-F\{q\}]+\psi R_p]}_{\text{Expected Benefit}} - \underbrace{(k+R_p)/\overline{\beta}}_{\text{Expected Cost}} \geq 0.$$

If $\overline{R}$ represents the R_p for which this condition is satisfied with equality, then we know that any p with $R_p < \overline{R}$ protests. Under the assumption that $R_p \sim U[0,1]$, $\Pr(R_p < \overline{R}) = \overline{R}$. If $\overline{R} = 0$, then no p protests. The government represses only if

$$c[1-(1-\psi)m[1-F\{q\}]-\psi\overline{R}] - R_G \geq c[1-\psi]$$
$$\underbrace{\psi(1-\overline{R})}_{\text{Deterred Protesters}} \geq \underbrace{(1-\psi)m[1-F\{q\}]}_{\text{Alienated Citizens}} + \frac{R_G}{c}.$$

□

B.3 Proof of Proposition 2 (Comparative Statics)

Proof. If the government does not repress, then an interest group member p protests if $c\psi \geq k/(\beta_0 + S\beta_s)$. This condition is more likely to hold as $S\beta_s$ increases.

If an interest group member p anticipates repression, then they protest if $c[(1-\psi)m[1-F\{q\}]+\psi R_p] \geq (k+R_p)/(\beta_0+S\beta_s)$. A protest only occurs if this condition is satisfied for the p with the smallest $R_p > 0$, and this condition is more likely to be satisfied for any p as $S\beta_s$ increases.

The government wants to repress if $\psi(1-\overline{R}) \geq (1-\psi)m[1-F\{q\}]+R_G/c$. This inequality is less likely to hold as m increases. The government never represses if this condition does not hold, regardless of whether protests actually take place.

Suppose that $c\psi \geq k/\overline{\beta}$ but $c[(1-\psi)m[1-F\{q\}]+\psi R_p] < (k+R_p)/\overline{\beta}$. If an increase in m shifts the government's decision from repression to no repression, then we move from a region in which no p protests to one in which all p protest. Thus, by disincentivizing repression, increasing m can increase the likelihood of protest. $\square$

C. Resolving Selection Problem for Repression Analysis

Estimating the effect of coverage on repression remains challenging. This is the case, because *repression is only observed when a protest actually takes place and not when a protest that would have been repressed never materializes* (i.e., when repression effectively deters protest).

Our theory helps reveal the thorniness of this selection problem, which can lead us to over- or under-estimate the true effect of cell phone coverage on the government's propensity to repress. Recall that our model has four equilibrium outcomes: (A) no protest, and government would not repress; (B) no protest, and government would repress; (C) protest, and government represses; and (D) protest, and government does not repress. If our argument is correct and cell phones reduce coordination costs and increase the visibility of repression, then receiving coverage can change the equilibrium in a locality in one of six ways. These are listed in the first column of table A.1.

Table A.1: The Selection Problem Related to Repression

Equilibrium Shift:			Actual Change:	Observed Change:	Proportion of
$D_i = 0$		$D_i = 1$	$\tau_i = R_i(1) - R_i(0)$	$\widetilde{\tau}_i = \widetilde{R}_i(1) - \widetilde{R}_i(0)$	Observations:
A	$\rightarrow$	B	1	0	p_{AB}
A	$\rightarrow$	C	1	1	p_{AC}
A	$\rightarrow$	D	0	0	p_{AD}
B	$\rightarrow$	C	0	1	p_{BC}
B	$\rightarrow$	D	-1	0	p_{BD}
C	$\rightarrow$	D	-1	-1	p_{CD}

(A) No protest, government would not repress; (B) No protest, government would repress; (C) Protest, government represses; (D) Protest, government does not repress.

How does true and observed use of repression change with each of these equilibrium shifts? Let $R_i(D_i)$ be the government's true decision about whether to employ repression in locality i as a function of i's treatment status, $D_i \in \{0, 1\}$. What we actually observe is $\widetilde{R}_i(D_i)$, which is one if a protest occurs in locality i and is repressed and zero otherwise. The second and third columns of table A.1 show the change in the true and observed use of repression, respectively. Taking the first row of the table as an example, when gaining cell phone coverage shifts an area from equilibrium A to equilibrium B, the government's decision to repress changes from 0 to 1 ($\tau = R(1) - R(0) =$

$1-0=1$), but we do not observe this change in repression because protest is deterred in equilibrium B ($\widetilde{\tau} = \widetilde{R}(1) - \widetilde{R}(0) = 0 - 0 = 0$). The final column of the table indicates the proportion that experience this equilibrium shift (e.g., p_{AB} is the proportion of localities that shift from A $\rightarrow$ B).

After weighting the actual change in repression by these proportions, the true average effect of cell phone coverage on repression can be written as:

$$\tau = p_{AB} + p_{AC} - p_{BD} - p_{CD}.$$

However, what we actually observe is:

$$\widetilde{\tau} = p_{AC} + p_{BC} - p_{CD}.$$

The true decrease in the use of repression is larger in magnitude than the observed reduction when the following condition holds:

$$\tau < \widetilde{\tau} \iff p_{AB} < p_{BC} + p_{BD}.$$

Put differently, when this condition holds, the selection problem makes it tougher to find evidence supporting our hypothesis that repression declines following the expansion of coverage.

This insight allows us to make *some* empirical progress. If we can remove the observations that make up p_{AB}, then, assuming our model is correct, our estimate *understates* the true reduction in repression that results from treatment. Equilibrium (A) (i.e., no protest, government would not repress) results when the costs protesting are high relative to the value of the concession, regardless of the government's response. In an attempt to exclude all such places, we drop localities that never experience a protest between 2000 and 2012 (or their first year of treatment, whichever comes first). Estimating equation 3 using the resulting sample, we feel more confident about interpreting the estimate of τ as understating the true reduction in repression that results from the introduction of cell phone coverage. This strategy allows us to plausibly recover a lower bound on the effect of coverage on repression.

D. Additional Empirical Analysis

D.1 Heterogeneous Effects by Regime

Table A.2: Coverage Expansion and Pr(Protest) by Regime; GDELT Data

	Dependent variable:		
	$\mathbb{1}$(Protest) × 100		
	(1)	(2)	(3)
$\mathbb{1}$(Covered)	0.088	0.001	−0.027
	(0.006)	(0.009)	(0.005)
$\mathbb{1}$(Free Media)		−0.026	
		(0.020)	
$\mathbb{1}$(Covered) × $\mathbb{1}$(Free Media)		0.335	
		(0.018)	
$\mathbb{1}$(Democracy)			0.075
			(0.016)
$\mathbb{1}$(Covered) × $\mathbb{1}$(Democracy)			0.225
			(0.011)
Cell FEs	2,110,209	2,101,218	2,063,528
Year FEs	6	4	6
Observations	12,661,254	8,352,108	12,081,395
Note:	Robust std. errors clustered on grid-cell.		

Notes: columns 1-3: linear probability model regressions, where the dependent variable has been multiplied by 100. See equation 1 for the econometric specifications and table 3 for information about the protest and coverage data. $\mathbb{1}$(Free Media) is based on the (country-year) Global Media Freedom data (Whitten-Woodring and Van Belle 2014). $\mathbb{1}$(Democracy) takes a one for country-years that receive a five or higher according to the Polity IV data (Marshall, Jaggers and Gurr 2012)

.

Table A.3: Indirect Effect of Coverage Expansion by Regime; GDELT Data

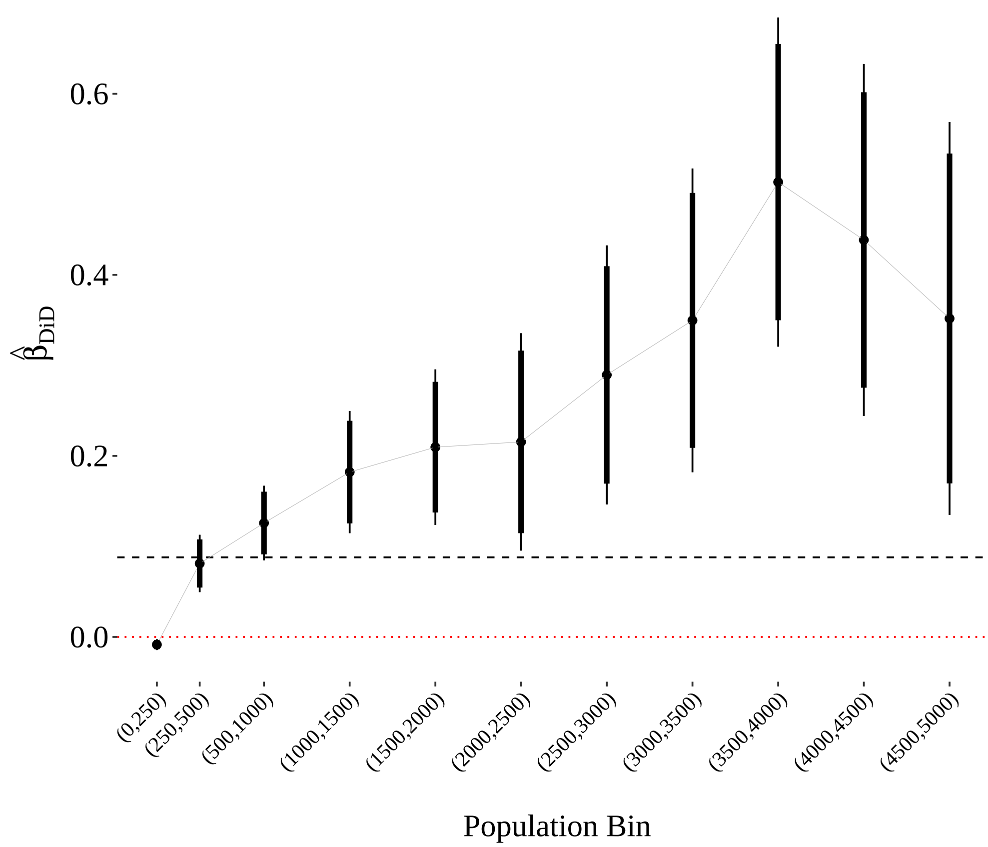
	Dependent variable:		
	$\mathbb{1}$(Protest) × 100		
	(1)	(2)	(3)
$\mathbb{1}$(Covered)	0.088	−0.027	0.061
	(0.006)	(0.005)	(0.057)
m			−0.046
			(0.019)
$\mathbb{1}$(Democracy)		0.075	−0.001
		(0.016)	(0.018)
$\mathbb{1}$(Covered) × m			−0.083
			(0.060)
$\mathbb{1}$(Covered) × $\mathbb{1}$(Democracy)		0.225	−0.508
		(0.011)	(0.100)
m × $\mathbb{1}$(Democracy)			0.147
			(0.025)
$\mathbb{1}$(Covered) × m × $\mathbb{1}$(Democracy)			0.764
			(0.106)
Cell FEs	2,110,209	2,063,528	2,063,528
Year FEs	6	6	6
Observations	12,661,254	12,081,395	12,081,395
Note:	Robust standard errors clustered on grid-cell.		

Notes: columns 1-3: linear probability model regressions, where the dependent variable has been multiplied by 100. See equation 2 for the econometric specifications and table 3 for information about the protest and coverage data. $\mathbb{1}$(Democracy) takes a one for country-years that receive a five or higher according to the Polity IV data (Marshall, Jaggers and Gurr 2012).

D.2 Heterogeneous Effects by Population

To understand whether urban, peri-urban, or rural areas are driving our main effects, we look at effect heterogeneity for grid cells with different levels of population (see figure A.2). We find that cell phone coverage has no effect in sparsely populated grid cells, where there are few residents that might coordinate and protest. These grid cells make up a meaningful proportion of our sample, attenuating our overall estimate.

Figure A.2: Coverage Expansion and Pr(Protest) by Population; GDELT Data



We bin cells into categories based on their population using LandSat data. We then interact our coverage indicator with those categories and plot the difference-in-differences estimates for cells with varying population sizes. We also include the 95% (and thicker 90%) confidence intervals. The dashed horizontal line is our estimate from table 3, model 1.

The effect size grows for more densely populated areas. (Our confidence intervals expand as there are fewer grid cells in these bins.) We note, however, that 5,000 inhabitants per grid cell

($\sim$140 inhabitants per km^2) corresponds to the population density of a town or small city, not major urban areas. If we look at grid cells with more than 36,000 inhabitants — a population density of 1,000 inhabitants per km^2 — the effect increases further to 1.97. We report this result and the linear interaction of our treatment indicator with population (divided by 1,000) in table A.4.

Table A.4: Coverage Expansion and Pr(Protest) by Population; GDELT Data

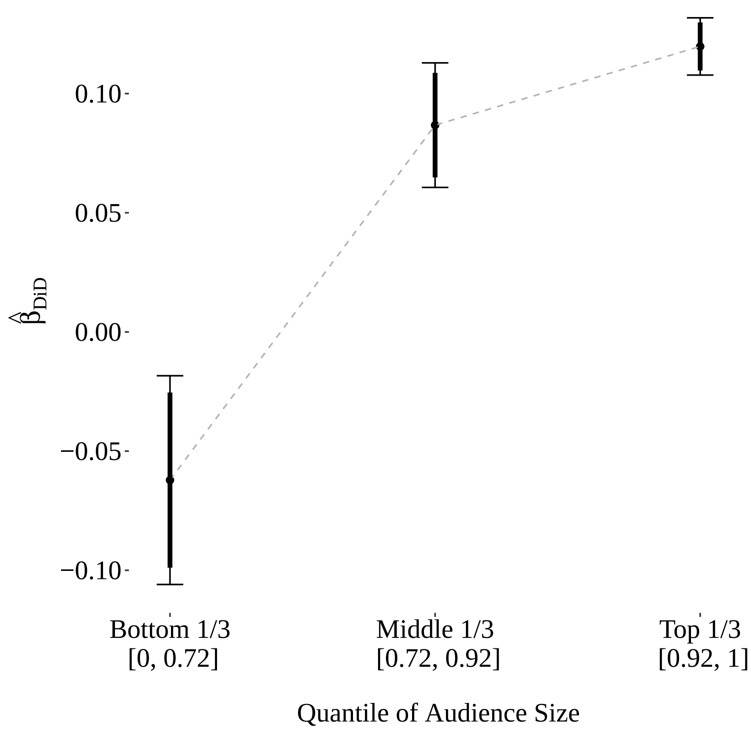
	Dependent variable:	
	$\mathbb{1}$(Protest) $\times$ 100	
	(1)	(2)
D_{it}	0.075	0.037
	(0.006)	(0.007)
$D_{it} \times \mathbb{1}$(Pop. $>$ 36k)	1.895	
	(0.287)	
$D_{it} \times$ (Pop./1000)		0.026
		(0.003)
Observations	12,661,254	12,661,254
Note:	Robust standard errors clustered on grid-cell.	

Notes: columns 1-2: linear probability model regressions, where the dependent variable has been multiplied by 100. See equation 1 for the econometric specifications and table 3 for information about the protest and coverage data. We use LandSat data to measure the population of each cell. Model 1 interacts our coverage indicator with an indicator for whether a cell exceeds 36,000 residents (or a density of roughly 1,000 / km^2); model 2 interacts the coverage indicator with a continuous measure of population (divided by 1,000).

D.3 Flexible Estimation of How the Effect of Coverage Varies with Audience Size (m)

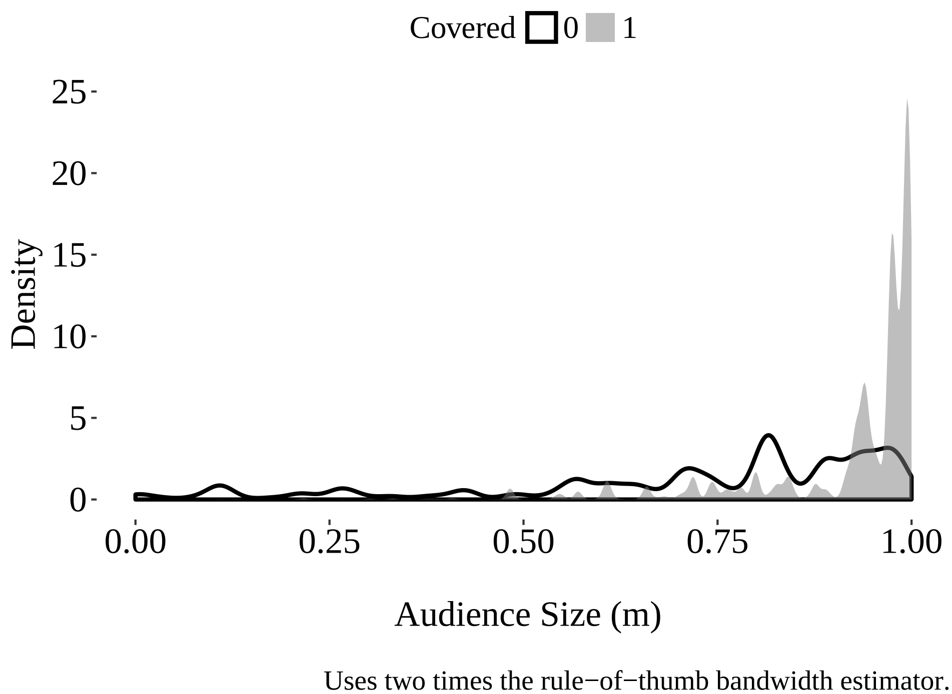
Rather than assuming that the conditional effect of coverage varies linearly with our measure of audience size (m), we estimate the effect of coverage at different quantiles of m. To do this, we amend equation 2, replacing our continuous measure of m with indicators for whether audience size falls in the bottom, middle, or top third of the sample. Figure A.3 shows that the effect of coverage increases substantially across these quantiles of m.

Figure A.3: Effect of Coverage at Different Quantiles of m



We amend equation 2 and interact our coverage indicator with the terciles of m, rather than the continuous measure. We then plot the estimated effect of cell phone coverage with 95% (and thicker 90%) confidence intervals for those different terciles, allowing a more flexible functional form.

Figure A.4: Distribution of m by Coverage



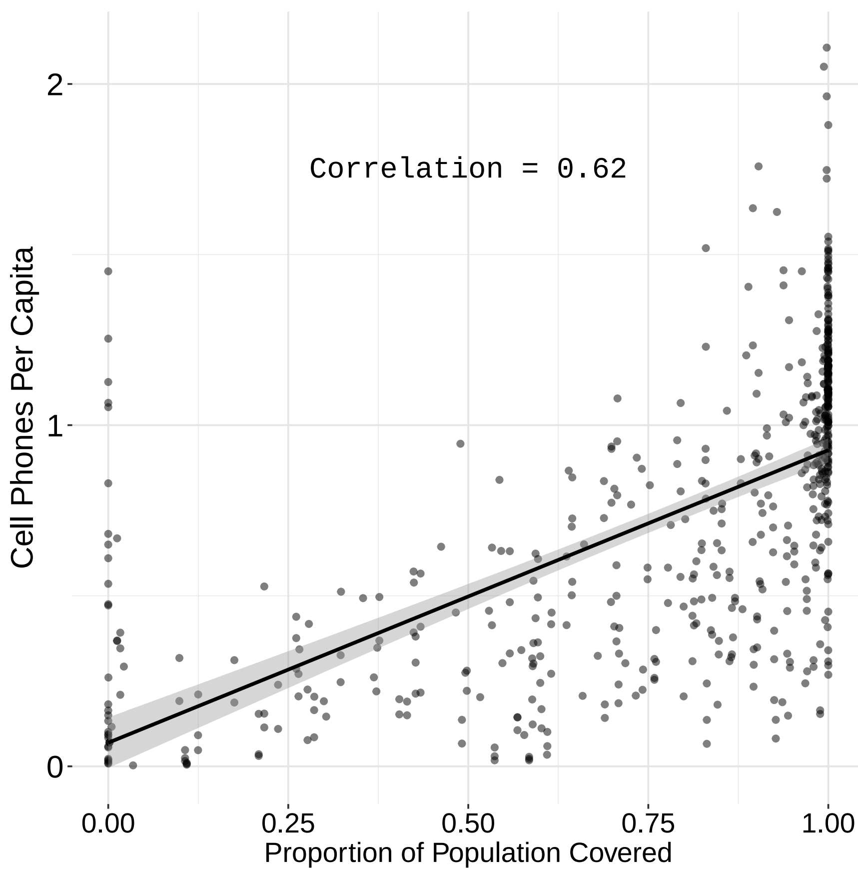
We plot the kernel densities of m for cells with (solid grey) and without (black outline) cell phone coverage. As is apparent from the figure, most covered cells occur in contexts where $m > 0.75$.

For reference, we also show the distribution of m for the sample used in table 3. As noted above, we caution against reading too much into the implied effect of coverage at very low-levels of m, as there are not many treated cells in this range.

D.4 Measure Validity: Comparing Cell Phone Coverage with Mobile Phone Ownership

The Collins Mobile Coverage Explorer database is compiled from submissions by telecom operators around the World. To check that reported expansions in coverage correspond to increases in cell phone use, figure A.5 compares the proportion of the population covered by the cell phone network (according to the Collins Mobile Coverage Explorer database) with data on cell phone ownership per capita from Banks and Wilson (2014). As expected, we find that the two are highly positively correlated ($\rho = 0.62$).

Figure A.5: Cell Phone Coverage vs. Cell Phone Ownership Per Capita



We calculate the proportion of the population covered by the cell phone network using the formula in section 4.3 and data from the Collins Mobile Coverage Explorer database and LandScan. Data on cell phone ownership per capita come from Banks and Wilson (2014). Note that cell phone ownership per capita can exceed one if the average individual owns multiple phones.

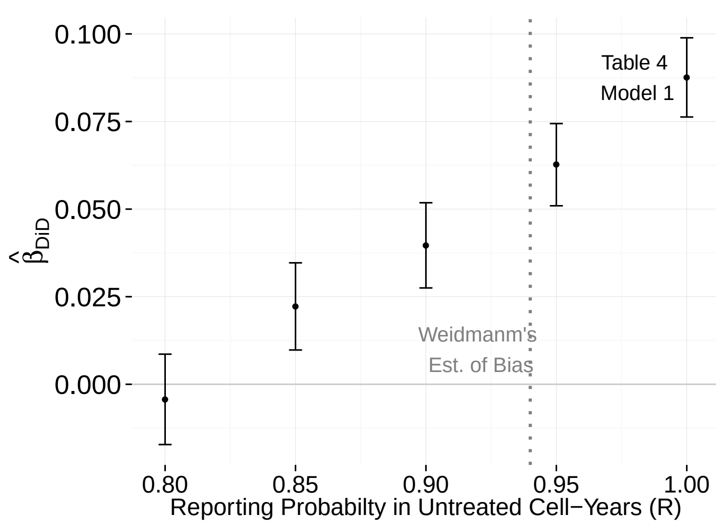
In some very small countries (e.g., the Bahamas, Djibouti, Kiribati) ownership is high despite minimal coverage. In particular, there are 40 country-years where the proportion of the population covered is less than 0.05, yet per capita ownership exceeds 0.25. This suggests that we may be wrongly classifying some areas as part of the control group when they, in fact, enjoy some access. Comfortingly, this works against rejecting the null. Furthermore, such observations make up less than 1% of our sample and, thus, do not meaningfully impact our results.

D.5 Bounding Reporting Bias as a Potential Confound

Figure A.6 indicates the the probability of reporting in treated and untreated areas would have to differ by more than 15 percentage points to explain away our effects — more than double the reporting bias that Weidmann (2015) estimates using data from Afghanistan. For the purposes of this bounding exercise, we assume that (1) there is no underreporting in treated areas; and (2) the null hypothesis of no difference in the probability of protest in areas with and without cell phone coverage. These assumptions imply that we can estimate the average probability of protest in all cells by just looking at treated cell-years. Call this probability $\mathcal{P} = \Pr(\text{Protest}|\text{Treated}) = \Pr(\text{Protest}|\text{Untreated})$. Let $\mathcal{R}$ be the probability that a protest is reported on if it occurs; our first assumption implies that $\mathcal{R} = 1$ in treated grid cells. With these assumptions, we proceed as follows:

- We retain the outcome information of treated grid cell-years.
- If a grid cell-year does not get coverage but reports a protest, we retain their outcome data.
- If a grid cell-year does not get coverage *and* does not report a protest, then we assume that a protest occurred with probability $\mathcal{P}$ and was reported on with probability $\mathcal{R}$. We thus assign new outcomes to these cells by drawing from $\{0, 1\}$ with probabilities $\{1 - \widehat{\mathcal{P}}\mathcal{R}, \widehat{\mathcal{P}}\mathcal{R}\}$.
- We use this new outcome vector to estimate equation 1 for different levels of reporting bias.

Figure A.6: Reporting Bias Required to Explain Away Our Results



Estimates assuming different levels of underreporting in uncovered areas relative to covered areas. Weidmann (2015) estimates this bias at 0.06 in Afghanistan (indicated with the dashed vertical line).

D.6 Robustness to Clustering on Larger Geographic Units

In the primary analysis, we cluster our standard errors on grid cell to account for temporal dependence. To account for possible spatial dependence, we also nest each of our 6×6 km cells in larger 24×24 km cells. Table A.5 replicates table 3 but clusters the standard errors on these larger (24×24 km) units. Our inferences are unchanged.

Table A.5: Coverage Expansion and Pr(Protest), Clustering on Larger Geographies; GDELT Data

	Dependent variable:				
	$\mathbb{1}$(Protest) $\times$ 100				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$(Covered)	0.088	0.037	-0.251	0.085	-0.237
	(0.006)	(0.006)	(0.055)	(0.006)	(0.055)
m			0.096		0.097
			(0.021)		(0.021)
$\mathbb{1}$(Covered) $\times m$			0.362		0.344
			(0.058)		(0.058)
Log Luminosity$_{t-1}$				0.033	0.028
				(0.007)	(0.007)
Cell FEs	2,110,209	2,110,209	2,110,209	2,110,209	2,110,209
Year FEs	6		6	6	6
Country$\times$Year FEs		1,236			
Observations	12,661,254	12,661,254	12,661,254	12,661,254	12,661,254

Note: Robust std. errors clustered on grid-cell.

D.7 Robustness to Using Cities as Unit of Analysis

In geo-coding events, GDELT assigns them to the town or city of occurrence. For this reason, our main analysis uses a grid with cells sized to correspond to the median city's area (6 × 6 km). We corroborate our results using a lower resolution (24 × 24 km). In this section, we also present a city-level analysis, in which the geographic units of analysis are contiguous areas with 200 people per km^2 or more according to Oak Ridge National Laboratory (2012). Of the 5,793 cities, our sample comprises the 927 cities that were not covered throughout the period of analysis. We code a city as covered by a cell phone network if any of its area is covered by a network in a given year. Results in table A.7 support our previous findings. When we include country-specific flexible time trends, we find that the direct effect of coverage is positive (if slightly smaller in magnitude). Moreover, we find strong evidence that the likelihod of protest increases as the size of the audience grows; at $m_{ct} = 0.78$ — which falls at the 17th percentile of covered cities — the effect becomes positive.

Table A.6: Summary Statistics

Statistic	N	Mean	St. Dev.
$\mathbb{1}$(Protest) × 100	5,562	17.080	37.640
$\mathbb{1}$(Covered)	5,562	0.401	0.490
m	5,562	0.677	0.320

Table A.7: City-Level Analysis

	Dependent variable:		
	$\mathbb{1}$(Protest) × 100		
	(1)	(2)	(3)
$\mathbb{1}$(Covered)	−0.129	3.962	−17.460
	(1.178)	(1.982)	(5.704)
m			−5.568
			(4.732)
$\mathbb{1}$(Covered) × m			22.500
			(6.420)
Cell FEs	927	927	927
Year FEs	6		6
Country×Year FEs		540	
Observations	5,562	5,562	5,562
Note:	Robust std. errors clustered on city.		

D.8 Effect of Coverage on Protest using Alternative Datasets

D.8.1 Results using ICEWS Data

We replicate our main analysis using the Integrated Crisis Early Warning System (**ICEWS**) data. ICEWS is a product of Lockheed Martin that draws on news sources from approximately 300 publishers, including international and national sources (Boschee et al. 2015). ICEWS machine codes events from this corpus using the Conflict and Mediation Event Observations (CAMEO) system, which includes a top-level category for protest (Schrodt and Yilmaz 2007). The dataset covers all countries over the period from 1995 to 2014. We limit the sample to events that name a specific city or town. A recent evaluation of the ICEWS data asked human coders to evaluate a sample of events (from 2011 to 2013) and determine (a) whether protest events were, in fact, protests, (b) whether the correct source actor was coded, and (c) whether the correct target actor was coded. The report found that 84.5% of protest events in the sample met these three criteria (Raytheon BBN Technologies 2015, 8).

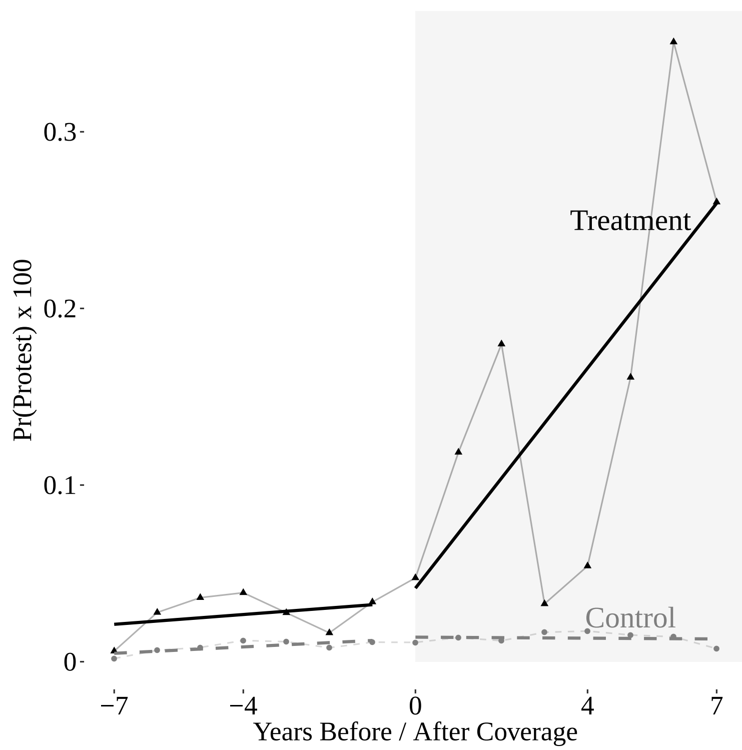
Table A.8: Pr(Protest) by Coverage; ICEWS Data

Never Covered	$\mathbb{1}$(Covered)	Pr(Protest) $\times$ 100 (GDELT)	Pr(Protest) $\times$ 100 (ICEWS)
1	0	0.050	0.007
0	0	0.186	0.030
0	1	0.459	0.062

Table A.9: Summary Statistics: ICEWS Data

Statistic	N	Mean	St. Dev.	Min	Max
$\mathbb{1}$(Protest) $\times$ 100	12,661,298	0.021	1.466	0	100
$\mathbb{1}$(Covered)	12,661,298	0.178	0.383	0	1

Figure A.7: Effect of Coverage Expansion on Pr(Protest); ICEWS Data
Trends in Pr(Protest) are parallel prior to treatment, but Pr(Protest) increases after cell phone coverage.



The figure plots the probability of protest in the years before and after coverage.

Table A.10: Coverage Expansion and Pr(Protest); ICEWS Data

	Dependent variable:				
	$\mathbb{1}$(Protest) $\times$ 100				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$(Covered)	0.004	0.002	−0.040	0.005	−0.041
	(0.002)	(0.003)	(0.030)	(0.002)	(0.031)
m			0.010		0.010
			(0.009)		(0.009)
$\mathbb{1}$(Covered) $\times$ m			0.047		0.049
			(0.032)		(0.032)
Log Luminosity$_{t-1}$				−0.002	−0.003
				(0.002)	(0.002)
Cell FEs	2,110,209	2,110,209	2,110,209	2,110,209	2,110,209
Year FEs	6		6	6	6
Country$\times$Year FEs		1,236			
Observations	12,661,298	12,661,298	12,661,298	12,661,298	12,661,298

Note: Robust std. errors clustered on grid-cell.

D.8.2 Results using SCAD Data

We estimate equation 1 using the Social Conflict in Africa Database (SCAD) described in section 4.2 (Hendrix and Salehyan 2012). This demonstrates, first, that our results are robust to the use of alternative, hand-coded data on protest events. Second, as the name indicates, the SCAD only covers African states; this alleviates concerns about changes in cellular conventions (e.g., to GSM from CDMA/IS-95 in the US) driving our findings. As a percentage of the sample mean of the dependent variable, these effects are roughly twice as large as those reported in table 3, model 1.

Table A.11: Pr(Protest) by Coverage; SCAD Data

Never Covered	$\mathbb{1}$(Covered)	Pr(Soc. Conf.) $\times$ 100	SD
1	0	0.009	0.946
0	0	0.006	0.743
0	1	0.025	1.579

Table A.12: Summary Statistics: SCAD Data

Statistic	N	Mean	St. Dev.	Min	Max
$\mathbb{1}$(Soc. Conf.) $\times$ 100	1,992,524	0.009	0.969	0	100
$\mathbb{1}$(Covered)	1,992,524	0.054	0.227	0	1
Log Luminosity$_{t-1}$	1,992,524	0.308	0.385	0.000	4.157

Table A.13: Coverage Expansion and Pr(Soc. Conf.); SCAD Data

	Dependent variable:		
	$\mathbb{1}$(Soc. Conf.) $\times$ 100		
	(1)	(2)	(3)
$\mathbb{1}$(Covered)	0.0189	0.0244	0.0189
	(0.0073)	(0.0094)	(0.0073)
Log Luminosity$_{t-1}$			-0.0008
			(0.0053)
Cell FEs	498,131	498,131	498,131
Year FEs	4		4
Country$\times$Year FEs		228	
Observations	1,992,524	1,992,524	1,992,524
Note:	Robust std. errors clustered on grid-cell; $^{\dagger}p < 0.1$, $^{*}p < 0.05$		

D.8.3 Agreement across GDELT, ICEWS, and SCAD

If over-reporting in the GDELT data is driven by cell-phone coverage, then we expect to see more discrepancies between GDELT and alternative protest datasets with treatment. However, table A.14 below indicates that there is no meaningful difference in the likelihood that GDELT and ICEWS or SCAD disagree after cells transition into coverage, increasing our confidence that cell coverage is not amplifying reporting bias in the GDELT data relative to the other event datasets.

Table A.14: Discrepancies across Protest Data by Cell Coverage

	Dependent variable:	
	$\mathbb{1}(\text{GDELT} \neq \text{ICEWS})$	$\mathbb{1}(\text{GDELT} \neq \text{SCAD})$
	(1)	(2)
$\mathbb{1}$(Covered)	0.00001	-0.0002
	(0.0001)	(0.0002)
Cell FEs	3,413,247	638,386
Year FEs	6	4
Observations	20,479,849	2,553,544
Note:	Robust std. errors clustered on grid-cell.	

D.9 Effect of Mobile Phone Ownership on Protest using Cross-National Data

While our high-resolution data allow us to employ a more credible empirical strategy than past work, our basic findings are not driven by our decision to focus on a much smaller unit of analysis (the grid cell) than is typical in cross-national comparative projects. In table A.15 we use the well-known Cross-National Time-Series Data Archive from Banks and Wilson (2014) to replicate our first result. Employing a country-year panel from 1991-2011, we find that cell phones per capita (lagged one year) are associated with a higher probability of protest and a higher number of protests (where protests include anti-government demonstrations, strikes, and riots). These models include country fixed effects and country-specific linear time trends, and controls for logged GDP and logged population.

Table A.15: Cross-national Correlations of Cell Phones (per capita) and Protest

	Dependent variable:			
	$\mathbb{1}$(Protest)		Number of Protests	
	(1)	(2)	(3)	(4)
Cell phones per capita (lag)	0.137	0.086	1.784	1.289
	(0.048)	(0.053)	(0.664)	(0.539)
Log of GDP per capita (lag)		0.020		0.418
		(0.032)		(0.290)
Log of population (lag)		0.123		1.598
		(0.231)		(4.540)
Country-specific trend	Yes	Yes	Yes	Yes
Country FEs	197	195	197	195
Observations	3,873	3,678	3,859	3,668
R^2	0.370	0.371	0.305	0.270
Note:	Robust std. errors clustered on country.			

Notes: columns 1-2: linear probability models. Columns 3-4: OLS regressions with the number of protests used as the dependent variable. Data for all variables is taken from the CNTS Data Archive from 1991-2011.

In all of the specifications, the correlations between cell phones per capita and our protest variables are positive; the relationship is statistically significant (or nearly significant) in all four models. The first model implies that one within-country standard deviation increase in cell phones per capita (0.16) is associated with a two percentage point increase in the probability of protest (or 14 percent of a within-country standard deviation of the dependent variable). While we are comforted by finding a similar correlation between cell phone penetration and protest activity at the country-level, this analysis is more likely confounded by omitted variables than our early results that leverage over time variation within very small geographic units. Furthermore, these country-level data does not allow us to evaluate our second hypothesis.

Supplemental Information References

Banks, Arthur S and Kenneth A Wilson. 2014. "Cross-National Time-Series Data Archive." Databanks International.

Boschee, Elizabeth, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz and Michael Ward. 2015. "ICEWS Coded Event Data." Harvard Dataverse.

Hendrix, Cullen S and Idean Salehyan. 2012. "SCAD 3.0 Codebook.".

Little, Andrew T. 2015. "Communication Technology and Protest." *The Journal of Politics* 78(1):152–166.

Marshall, Monty, Keith Jaggers and Ted Gurr. 2012. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010.".

Morris, Stephen and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review* 92(5):1521–1534.

Oak Ridge National Laboratory. 2012. "LandScan." Oak Ridge National Laboratory.

Raytheon BBN Technologies. 2015. BBN ACCENT Event Coding Evaluation. Technical report.

Schrodt, Philip A and Omur Yilmaz. 2007. "CAMEO Conflict and Mediation Event Observations Codebook.".

Weidmann, Nils. 2015. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* pp. 1–13.

Whitten-Woodring, Jenifer and Douglas A Van Belle. 2014. *Historical Guide to World Media Freedom: A Country-by-Country Analysis*. Sage/CQ Press.