

SUNLIGHT DISINFECTS? FREE MEDIA IN WEAK DEMOCRACIES

Leopoldo Fergusson
Juan Fernando Vargas
Mauricio A. Vela

SERIE DOCUMENTOS DE TRABAJO

No. 132

Marzo de 2018

Sunlight Disinfects?

Free Media in Weak Democracies*

Leopoldo Fergusson[†]

Juan F. Vargas[‡]

Mauricio A. Vela[§]

March 14, 2018

Abstract

Free media may not favor political accountability when other democratic institutions are weak, and may even bring undesirable unintended consequences. We propose a simple model in which candidates running for office may engage in coercion to obtain votes. A media scandal exposing these candidates entices them to increase their coercion effort to offset the negative shock on their popularity, potentially minimizing or even counteracting the effect of the scandal on their vote share. We provide empirical evidence from one recent episode in the political history of Colombia in which politicians seeking a seat in Congress colluded with illegal armed paramilitary groups to obtain votes, and this collusion was ultimately brought to light by a media scandal. We find that paramilitary-backed candidates exposed before elections get as many votes as those exposed after elections, but their electoral support is more strongly concentrated where coercion is cheaper: in areas with paramilitary presence and weak state capacity. Our results highlight the complementarity between different dimensions of democratic institutions.

Keywords: Media, Democracy, Elections, Colombia, Civil Conflict, Coercion.

JEL: D72, D74, L82, P16.

*This is a substantially revised version of a working paper circulated under the same title in 2013. We thank comments from Daron Acemoglu, Ruben Durante, Claudio Ferraz, Frederico Finan, Claudia López, Abbey Steele, Eik Swée, and seminar participants at Universidad de los Andes, Universidad del Rosario, Banco de la República, LACEA-LAMES 2017 Annual Conference, the Xth Workshop on Media Economics, the 8th Annual Households in Conflict Network Workshop, and MIT's Political Economy Breakfast. We also thank Fidel Cano, Pablo Farfán and Carolina Reyes at *El Espectador*, Mónica Pachón for help with data on congressmen characteristics, and Daron Acemoglu, Rafael Santos, and James Robinson for sharing their measures on paramilitary and guerrilla presence. Financial support from MIT's George and Obie Shultz Fund and Universidad del Rosario is gratefully acknowledged. We also thank Mariana Blanco and Patricio Dalton for sponsoring part of our data collection effort.

[†]Universidad de los Andes, Facultad de Economía, Cra 1 N° 18A - 12 Bogotá, Colombia. Phone: +57 1 3394949, ext. 2439. E-mail: lfergusson@uniandes.edu.co

[‡]Universidad del Rosario, Facultad de Economía, Cl 14 N° 4 - 80 Bogotá Colombia. Phone: + 57 1 2970200, ext. 4129. E-mail: juan.vargas@urosario.edu.co

[§]Washington University in St. Louis, Department of Political Science, One Brookings Drive, St. Louis, MO 63130-4899. Phone: +1 (703) 5995414. E-mail: mauriciovelabaron@go.wustl.edu

1 Introduction

A free and active media is recognized as essential for political accountability. By providing information, mass media can help voters make better decisions and hold politicians accountable. Often, journalists also help uncover corruption scandals and undue influence of special interests' groups. A famous example from US history is the Progressive Era, when the press mobilized the population against the power and abuses of Robber Barons and their trusts.¹ It is precisely around this time when Louis Brandeis, who championed the Progressive ideals, famously remarked: "Sunlight is said to be the best of disinfectants".² Earlier, Thomas Jefferson went so far as to say that free media is a sufficient condition for political accountability: "The functionaries of every government have propensities to command at will the liberty and property of their constituents. There is no safe deposit for these but with the people themselves, nor can they be safe with them without information. *Where the press is free, and every man able to read, all is safe.*" (Letter to Charles Yancey, 1816. ME 14:384, quoted in Snyder and Strömberg (2010), emphasis added).

In this paper, we argue that Jefferson was wrong: free media *per se* is no guarantee of political accountability. Unless free media operate in a sufficiently strong institutional environment, provision of information about abuse or misbehavior may not increase political accountability, and might even have unintended negative consequences. In particular, in our setting the exposure of misdeeds by politicians seeking office can aggravate electoral manipulation (where they have the option) to make up for a decline in electoral support elsewhere.

To discipline these ideas, we propose a simple political economy model where two types of politicians compete in a multiple-district election. One of the politicians may coerce a fraction of voters into voting for him. This politician must decide both the policy platform to offer to voters and the level of coercion to exert.

In this context, we study the implications of media exposure. We assume that media exposure of a politician's wrongdoings decreases his popularity relative to his opponent. While this effect reduces electoral support for the exposed politician across all districts, there are now more incentives to invest in coercion. Hence, the media scandal may *increase* the vote share of the exposed politician in places where, by increasing coercion, he fully compensates for the decrease in popularity due to exposure. We show that this is more likely to occur where the media scandal is intense, where voters are less ideologically biased towards either party, and where local institutions are weak. Each of these parameters increase the politician's incentives to coerce following the media scandal: a more intense scandal and less ideological voters imply that more votes are lost as a result of the scandal,

¹Glaeser, Gentzkow, and Goldin (2004) argue that the expansion of the US newspaper industry between 1870 and 1920 partly caused the decline of political corruption in that period. Gentzkow, Shapiro, and Sinkinson (2011) show that the entry of US daily newspapers from 1869-1928 produced a more active electorate, increasing voter turnout. Acemoglu and Robinson (2012) highlight the role of the free press in exposing the excesses of Robber Barons, and other scandals of political corruption at the local and federal levels.

²See <http://www.law.louisville.edu/library/collections/brandeis/node/196>.

and stronger institutions increase the marginal cost of coercion.

Beyond the increase in coercion, the model shows that the media scandal has additional negative unintended consequences that derive from this reaction. First, since votes may increase in some districts where coercion is strong enough, the scandal might not hurt the overall electoral success of the exposed politician, undermining the accountability benefits of exposure.³ Second, both politicians only consider the preferences of non-coerced voters, who unlike coerced voters respond to policy proposals when deciding whom to vote for. Thus, increased coercion after the scandal implies that politicians pay less attention to voters' needs when selecting policies.

This highlights the complementarity between different dimensions of democratic institutions, as having a free and active media may not be enough and may even bring undesirable unintended consequences if other dimensions of institutional quality, like the existence of free and fair elections, are not guaranteed.

We test some of the implications of our model in the context of Colombia's legislative elections, using one salient episode of the period from 2002 to 2010: the 'parapolitics' scandal that brought to light the alliance between politicians and illegal armed paramilitary groups to obtain votes by coercive means in exchange of a lenient legislation toward these criminals (Acemoglu, Robinson, & Santos, 2013). The timing and scope of the media scandal is revealed in Figure 1, which shows a surge in the share of parapolitics news stories. While the media scandal had positive effects in that it certainly led to judicial investigations of (and jail sentences to) many of the tainted politicians, we show that it also had the type of unintended negative consequences we highlight in our model.⁴

We cannot capture variation in coercion reliably, given the inherent difficulty in measuring any illegal activity. However, since we have good measures at the municipality level of the factors that make coercion more (state presence) or less (paramilitary presence) costly, we can verify whether the voting patterns in the data are consistent with the model.⁵ We first document that 'parapoliticians' have a different vote distribution than competitors: Senate candidates involved in the scandal get significantly more votes in areas with paramilitary presence and less state capacity. Most importantly, we directly test the reaction to the scandal by comparing, within parapoliticians, candidates exposed by the press before elections with those exposed after elections. We also verify that these candidates share similar individual characteristics, to help rule out the concern that those exposed late are fundamentally different and do not constitute a valid control group for the ones exposed early. Consistent with the increase in coercion predicted by the model, parapoliticians exposed before the elections shift their distribution of votes to areas where coercion is easier to exert. These results are robust to including candidate-level controls, municipal fixed effects and election-

³The wellbeing of exposed politicians, on the other hand, falls with media exposure because coercion is costly.

⁴For a description of how the media helped avoid impunity in parapolitics cases, see Restrepo (2008).

⁵"Coercion" is used here broadly to capture various form of electoral manipulation, which in our empirical application included the intimidation of voters through terror, massacres and political assassinations, as well as electorally fraudulent practices like vote buying, ballot stuffing and the use of the voting IDs of deceased people (Valencia, 2007).

year fixed effects. Finally, we provide evidence that parapoliticians are relatively more successful than their clean competitors in terms of their vote share. Also, relative to other parapoliticians, those involved in the scandal before elections were able to increase electoral support where coercion is more likely, and sufficiently so that their vote share does not significantly differ from that of candidates exposed after the elections.

To complement our main findings, we discuss additional anecdotal evidence suggesting that coercion is the likely underlying mechanism explaining these voting patterns. We also find that, among parapoliticians, those exposed by the media before elections get more votes from places with unusually high levels of turnout than those exposed after elections. These results are also in line with coercion being the underlying mechanism of our main results, as most forms of coercion by paramilitaries documented in the media and judicial investigations inflate vote totals.

Though the role of media in democracy has long been studied, our paper relates most closely to the relatively recent economics research on the behavior and political effects of mass media (surveyed by Prat and Strömberg (2013)). According to this line of research, access to media and its content influences citizens' electoral behavior and attitudes, increasing turnout (Gentzkow et al., 2011), improving information about representatives (Snyder & Strömberg, 2010), and influencing vote shares for certain candidates (DellaVigna & Kaplan, 2007; Prat & Strömberg, 2005). Moreover, the influence of mass media extends to policy by empowering citizens (Besley & Burgess, 2002; Strömberg, 2004a, 2004b), and by making some issues more salient than others (Eisenstein & Strömberg, 2007).

A key message from the literature is that, by improving information and electoral response, mass media may improve political accountability (see, for example, Ferraz and Finan (2008), Snyder and Strömberg (2010) and Brunetti and Weder (2003)). Nevertheless, the power of the media is a double-edged sword. For instance, mass media can present biased information, its bias may persist over time, and it may influence electoral behavior (Baron, 2006; Durante & Knight, 2012; Gentzkow & Shapiro, 2004, 2006, 2008, 2010; Groseclose & Milyo, 2005; Puglisi, 2011). Along the same lines, while access to free media may be key for democratic consolidation (Andriantsoa et al., 2005; Glaeser et al., 2004) and even development (Djankov & McLiesh, 2002), with media "captured" by the government or other powerful actors, checks and balances may be dismantled rather than strengthened (McMillan & Zoido, 2002; Besley & Prat, 2006; Petrova, 2008; Adena, Enikolopov, Petrova, Santarosa, & Zhuravskaya, 2015).

Our paper contributes to this literature by emphasizing the limits of mass media in achieving political accountability, and proposing a novel mechanism. In particular, instead of highlighting the potentially negative effects of mass media bias due to political or economic capture, or the risks where an active media is absent because of political oppression or other causes, we argue that even if mass media is independent and provides valuable information to voters, in a weakly-institutionalized environment it may not promote political accountability and have unintended

negative consequences given the response of powerful affected actors.

A few recent papers also emphasize unintended consequences of information and media exposure, but the mechanisms highlighted differ from ours. For example, DellaVigna, Enikolopov, Mironova, Petrova, and Zhuravskaya (2011) show that media may have an unintended effect of increasing ethnic animosity, Chong, De La O, Karlan, and Wantchekon (2011) provide experimental evidence that information about corruption in Mexico may be insufficient to improve political accountability because discouraged voters partly respond by withdrawing from the political process, and Jetter (2017) shows that media attention to terrorism may encourage further attacks.

We also contribute to the literature on electoral fraud and violent electoral coercion (Callen & Long, 2015; Chaves, Fergusson, & Robinson, 2015; Asunka, Brierley, Golden, Kramon, & Ofosu, 2017). In particular, our finding that coercion decreases with the average ideological support for the parapolitician is consistent with Robinson and Torvik (2009) and Collier and Vicente (2012), who argue that violence is targeted toward ‘swing voters’. However, in contrast to Collier and Vicente (2014) (but consistent with Acemoglu et al. (2013) and with the Colombian historical account), in our model higher turnout is associated with more violent coercion, not with less. Also, we find that violent coercion is used to shape the outcome of elections. While this is consistent with authors such as Harish and Little (2017), this is not the case in other contexts. For instance Daniele and Dipoppa (2017) document that mafia-like criminal organizations exert violence after the election, to influence the behavior of elected politicians.

Finally, our paper also relates to the idea that political elites often exercise *de facto* power to maintain their political power and shape institutions (Acemoglu & Robinson, 2008). The empirical literature provides some evidence along these lines (e.g. Naidu, 2012; Fergusson, Querubin, Ruiz-Guarin, & Vargas, 2017). However, we still do not fully understand the repercussions that these endogenous reactions by powerful elites have on society. Our results suggests that there may be negative unintended consequences beyond the mere persistent of political power of organized elites.

We next present our model in section 2. We derive testable empirical implications and discuss anecdotal evidence consistent with them. Section 3 presents our main data and section 4 discusses our empirical strategy to examine the theoretical implications more systematically. In section 5 we discuss our main results and robustness checks, and section 6 concludes. The Appendix discusses details on data construction, an extension to the baseline model, and additional robustness checks.

2 A simple model

To guide our empirical analysis, in this section we present a model of the electoral consequences of media exposure of parapoliticians. We consider an environment with $m = 1, \dots, N$ municipalities, each with a mass of voters of size 1. Two politicians, A and B, run for office, but A (the parapolitician) is able to collude with illegal armed groups (paramilitaries) that force voters to vote in a

particular way. As in Acemoglu et al. (2013), we assume that when coercion occurs, the population votes for politician A. However, in contrast to them, we assume that only a share $\lambda_m \leq 1$ of the population is coerced and votes for A. The remaining fraction $1 - \lambda_m$ of voters freely decide whom to vote for and we model this decision as in the standard probabilistic voting model (Lindbeck & Weibull, 1987). More specifically, a free voter i in municipality m votes for A so long as:

$$u^m(\mathbf{q}^A) + \sigma^{im} > u^m(\mathbf{q}^B),$$

where \mathbf{q}^x is the policy vector offered by politician $x \in \{A, B\}$, $u^m(\cdot)$ is the indirect utility over policies for citizens in municipality m , and σ^{im} is the ideological bias of voter i in municipality m in favor of politician A relative to politician B. We assume σ^{im} is distributed uniformly with mean $\sigma_m - \mu_m$ and density ϕ_m (hence, over the support $\left[-\frac{1}{2\phi_m} + \sigma_m - \mu_m, \frac{1}{2\phi_m} + \sigma_m - \mu_m\right]$). The mean of σ^{im} in each municipality has two components. First, σ_m is the average relative ideological leaning toward party A absent a media scandal in municipality m . Second, if a media scandal breaks out and newspapers publish the collusion of illegal armed groups with party A, the ideological bias in favor of A falls across all municipalities. This is captured by μ_m , measuring a potentially different impact of the media scandal in each municipality m . We also allow ϕ_m to vary across municipalities, with a larger ϕ_m meaning less ideological dispersion and therefore stronger response of voters to policy proposals.

Finally, the cost of coercion is given by

$$C(\{\lambda_m\}_{m=1}^N) = \sum_{m=1}^N \psi_m c(\lambda_m),$$

with increasing and convex costs, $c' > 0$, $c'' > 0$. To keep our model as simple as possible, we assume a quadratic cost, $c(\lambda_m) = \lambda_m^2/2$. We interpret the parameter ψ_m as local state capacity, which increases the cost of coercion.

To solve this simple game, we find the policy positions \mathbf{q}^A and \mathbf{q}^B that politicians A and B simultaneously offer to maximize their total electoral support⁶. Policy proposals are credible (politicians fulfill their promises). Politician A, who has colluded with paramilitaries, also chooses the level of coercion effort λ_m to exert in each municipality. A more realistic formulation would incorporate an exchange between politician A and paramilitaries, who in turn select coercion. However, since we are not directly interested in the game between politicians and paramilitaries, we simplify by assuming that A selects λ_m directly.

⁶The implications of the model when assuming that politicians maximize their probability of winning are similar. However, in the context of the Senate elections we analyze (open or closed list proportional representation) it seems more natural to model each politician as maximizing his votes. Our assumption of two candidates running for office is just for analytical simplicity.

2.1 Equilibrium policies and coercion

To find the equilibrium, we calculate the share of votes for each politician as a function of policy proposals. First we find the share of individuals who vote freely that support A in each municipality m using the distribution of σ^{im} . More specifically, we find the fraction of free voters for whom $\sigma^{im} > u^m(\mathbf{q}^B) - u^m(\mathbf{q}^A)$:

$$\frac{1}{2} + \phi_m (\sigma_m - \mu_m + u^m(\mathbf{q}^A) - u^m(\mathbf{q}^B)).$$

Since in each municipality a share λ_m is forced to vote for A, total support for A in each municipality, π_m^A , is

$$\begin{aligned} \pi_m^A &= \lambda_m + (1 - \lambda_m) \left[\frac{1}{2} + \phi_m (\sigma_m - \mu_m + u^m(\mathbf{q}^A) - u^m(\mathbf{q}^B)) \right] \\ &= \frac{1 + \lambda_m}{2} + (1 - \lambda_m) \phi_m (\sigma_m - \mu_m + u^m(\mathbf{q}^A) - u^m(\mathbf{q}^B)). \end{aligned} \quad (1)$$

Adding over all municipalities, total electoral support for A, Π^A , is:

$$\Pi^A = \sum_m \pi_m^A = \frac{N + \sum_m \lambda_m}{2} + \sum_m \phi_m (1 - \lambda_m) (\sigma_m - \mu_m + u^m(\mathbf{q}^A) - u^m(\mathbf{q}^B)).$$

Politician A chooses policies and coercion to maximize $\Pi^A - C(\{\lambda_m\}_{m=1}^N)$. In turn, politician B chooses policies to maximize $\Pi^B = N - \Pi^A$. The solution (assumed interior) to this problem is policy convergence with \mathbf{q}^* that satisfies:

$$\sum_{m=1}^N (1 - \lambda_m^*) \phi_m \nabla u^m(\mathbf{q}^*) = \mathbf{0}. \quad (2)$$

From the first order conditions of politician A, equilibrium coercion λ_m^* can be written as:

$$\lambda_m^* = \max \left\{ 0, \frac{1}{\psi_m} \left(\frac{1}{2} + \phi_m (\mu_m - \sigma_m) \right) \right\}, \quad (3)$$

for each municipality m . These results and their comparative statics implications are summarized in the following proposition.

Proposition 1 *Equilibrium policies and coercion*

In the equilibrium of the electoral game:

1. Parties A and B converge to the vector of policy proposals \mathbf{q}^* in (2). The resulting policy attention to voters in municipality m is decreasing in coercion λ_m^* and increasing in the density of ideological biases ϕ_m .

2. Party A exerts coercion λ_m^* in each municipality as given by (3). λ_m^* satisfies the following comparative static properties:

$$\begin{aligned}\frac{\partial \lambda_m^*}{\partial \psi_m} &\leq 0, \quad \frac{\partial \lambda_m^*}{\partial \sigma_m} \leq 0, \\ \frac{\partial \lambda_m^*}{\partial \mu_m} &\geq 0, \\ \frac{\partial \lambda_m^*}{\partial \phi_m} &\gtrless 0 \quad \text{if } \sigma_m - \mu_m \lesseqgtr 0.\end{aligned}$$

Proof. Follows directly from (2) and (3), derived in the main text. ■

Proposition 1 emphasizes that, as in any probabilistic voting model, municipalities where the ideological biases σ^{im} of the population are less densely distributed (that is, where ϕ_m is small) receive less policy attention. Indeed, with small ϕ_m there are more voters with strong biases towards party A or party B. Such voters will pay little attention to policy proposals in their voting decision. Thus, politicians have little incentive to cater to the policy needs of such voters.

In addition, and more important for our analysis, the utility of voters in each municipality receives a weight that is decreasing in the level of coercion. With coercion, only a share $1 - \lambda_m^*$ of citizens react to policy. Thus, both politicians take more into account the preferences of citizens in municipalities where there is less coercion.⁷ A useful way of thinking about the solution in (2) is to notice that it is equivalent to the solution of maximizing a social welfare \mathcal{W} function that gives a weight to citizens of municipality m that is proportional to $1 - \lambda_m^*$,

$$\mathcal{W} = \sum_{m=1}^N (1 - \lambda_m^*) \phi_m u^m(q).$$

Hence, even ignoring differences in voters' ideological biases (if $\phi_m = \phi$ for all m) only in municipalities without coercion ($\lambda_m^* = 0$) citizens' utility is fully taken into account in the political equilibrium. In this sense, coercion introduces a distortion relative to the choice of a social planner maximizing the sum of citizens' utilities.

Turning to equilibrium coercion, λ_m^* is decreasing in the local state capacity ψ_m , as better institutions make coercion more costly. On the other hand, coercion increases with the impact of the media scandal (μ_m): when politician A's popularity is adversely affected by the media, he tries to counteract the effect on electoral support by coercing more voters. Finally, notice that where

⁷This is similar to the Acemoglu et al. (2013) case of a solution with passive paramilitaries, only that they assume that no voters are able to respond to policy in the subset of paramilitary-controlled areas where coercion occurs. In our model, such set of municipalities correspond to those for which $\frac{1}{2} + \phi_m(\mu_m - \sigma_m) < 0$. One can directly verify that, given the distribution of ideological biases, in such municipalities all voters would freely vote for A ($\pi_m^A = 1$) and there is no incentive to coerce. In other words, given our simple formulation where the marginal cost of exerting minimal coercion approaches zero, parapoliticians exert at least some coercion in any municipality with positive support for the opponent B.

the population is more ideologically biased in favor of A, there is less coercion. Indeed, in such case coercion is not as necessary, as most voters will vote for A, driven solely by their ideological sympathy. This observation is also important to understand why the impact on coercion of the density ϕ_m of ideological biases depends on the sign of $\sigma_m - \mu_m$. Since μ_m measures the fall in average popularity of A following the media scandal and σ_m is the average ideological bias in favor of A absent any media scandal, $\sigma_m - \mu_m$ measures the average *net* relative sympathy towards A in the population, after the scandal. When this is negative, given policy convergence, the majority of voters is inclined to vote against party A. Moreover, the mass of voters that will favor B instead is proportional to ϕ_m . For this reason, when $\sigma_m - \mu_m < 0$ and ϕ_m increases, to avoid losing this electoral support party A responds by increasing coercion.

2.2 Support for the parapolitician

With policies and coercion described in Proposition 1, we can find the equilibrium share of votes received by party A in each municipality m :

$$\pi_m^A(\mu_m, \lambda_m^*) = \frac{1 + \lambda_m^*}{2} + (1 - \lambda_m^*)\phi_m(\sigma_m - \mu_m). \quad (4)$$

Adding over all municipalities, total equilibrium support for A is

$$\Pi^A = \frac{N + \sum_m \lambda_m^*}{2} + \sum_m \phi_m(1 - \lambda_m^*)(\sigma_m - \mu_m). \quad (5)$$

The expression highlights that party A, all else equal, has more votes than party B because paramilitaries give him coerced votes ($\sum \lambda_m^*/2 > 0$). More importantly, the impact of the media scandal on the electoral support for the parapolitician combines two effects that appear in the following expression for the derivative of $\pi_m^A(\mu_m, \lambda_m^*)$ with respect to μ_m :

$$\begin{aligned} \frac{d\pi_m^A}{d\mu_m} &= \frac{\partial \pi_m^A}{\partial \mu_m} + \frac{\partial \pi_m^A}{\partial \lambda_m^*} \frac{\partial \lambda_m^*}{\partial \mu_m} \\ &= -(1 - \lambda_m^*)\phi_m + \left(\frac{1}{2} + \phi_m(\mu_m - \sigma_m) \right) \frac{\partial \lambda_m^*}{\partial \mu_m} \geq 0. \end{aligned} \quad (6)$$

On the one hand, holding coercion fixed, support for A falls because the media scandal makes him less popular. This is the first term in (6), capturing the *direct* effect of the media scandal on parapolitician support. On the other hand, as the second term shows, as long as there is positive coercion in municipality m , there is an *indirect* effect whereby A at least partially compensates for the fall in free votes by endogenously increasing coerced votes⁸.

⁸That the indirect effect is always nonnegative follows from noting in our expression for coercion (3) that coercion λ_m^* is positive and $\partial \lambda_m^* / \partial \mu_m > 0$ whenever $1/2 + \phi_m(\mu_m - \sigma_m) > 0$.

Let's examine more closely how the media scandal affects each of the two terms in the indirect effect, namely: the rise in coercion following the media scandal ($\partial\lambda_m^*/\partial\mu_m$), and the electoral return of such increase in coercion ($\partial\pi_m^A/\partial\lambda_m$). The key observation is that both of these terms are increasing functions of the media scandal. Hence, the indirect effect which is the product of these two may dominate the direct effect under some parameter values (and, specifically, when the media scandal is large enough).

Starting with the electoral returns of an increase in coercion, note that this is simply

$$\partial\pi_m^A/\partial\lambda_m = 1 - (1/2 + \phi_m(\sigma_m - \mu_m)).$$

The first term in this expression tells us that an increase in coercion guarantees getting an additional vote in municipality m . However, since some of the free voters support A , there is no increase in support due to coercion from those $1/2 + \phi_m(\sigma_m - \mu_m)$ free votes. The difference is the *net* gain in votes with a marginal increase in coercion appearing in (6), and is increasing in the media scandal μ_m . Since a media scandal makes the parapolitician less popular, coercion is more important when less free voters are supporting him. This last observation also explains why the other term in the indirect effect, $\partial\lambda_m^*/\partial\mu_m$, is positive as noted already in Proposition 1.

In short, both coercion and the return to each unit of additional coercion increase when μ_m increases, implying that under some conditions the media scandal may increase the total electoral support for the parapolitician.

Finally, suppose that to capture the impact of a national media scandal, parametrized by μ , we can write the local impact of this scandal simply as $\mu_m = \mu l_m$ where $l_m \in [0, 1]$ measures the quality of local media and other factors influencing local diffusion. Then,

$$\frac{\partial\Pi^A}{\partial\mu} = \sum_{m=1}^N \frac{d\pi_m^A}{d\mu_m} \frac{\partial\mu_m}{\partial\mu} = \sum_{m=1}^N \frac{d\pi_m^A}{d\mu_m} l_m, \quad (7)$$

where $\partial\pi_m^A/\partial\mu_m$ is given by (6). These results, and other features of the electoral response to a media scandal, are summarized in the next proposition.

Proposition 2 *Equilibrium support for the parapolitician*

In the equilibrium of the electoral game, total votes for A (the parapolitician) in each municipality m are given by π_m^A in (4), and at the national level are given by Π_A in (25).

Moreover, $d\pi_m^A/d\mu_m$ is given by (6) and satisfies the following property:

$$\frac{d\pi_m^A}{d\mu_m} \geq 0 \quad \text{if} \quad \lambda_m^* \geq \frac{1}{2} \quad \left(\text{or} \quad \frac{1}{2} + \phi_m(\mu_m - \sigma_m) \geq \frac{1}{2}\psi_m \right)$$

Finally, the impact of a national media scandal μ on total support for the parapolitician, where

$\mu_m = \mu l_m$ with $l_m \in [0, 1]$, is given by:

$$\frac{\partial \Pi^A}{\partial \mu} = \sum_{m=1}^N \frac{d\pi_m^A}{d\mu_m} l_m, \gtrless 0$$

Proof. All results were derived in the main text, except for the conditions under which $d\pi_m^A/d\mu_m \gtrless 0$. This follows from using λ^* in (3) to write (6) as

$$\frac{d\pi_m^A}{d\mu_m} = -(1 - \lambda_m^*)\phi_m + \frac{\partial \lambda_m^*}{\partial \mu_m} \lambda_m^* \psi_m \gtrless 0.$$

From (3), we also have $\partial \lambda_m^* / \partial \mu_m = \phi_m / \psi_m$. After substituting, we get $d\pi_m^A/d\mu_m = -(1 - 2\lambda_m^*)\phi_m$ and the expressions in the proposition follow directly. ■

The most important feature of the solution is that, with the endogenous response of parapoliticians, the media scandal need not reduce their vote share where they can exert coercion. The media scandal may even *increase* parapolitician vote share in coerced municipalities when the direct negative impact of the scandal is more than compensated by the endogenous response in coercion effort from parapoliticians in collusion with paramilitaries. Proposition 2 shows that this occurs as long ϕ_m and μ_m are large enough. Each of these parameters increase parapoliticians' incentives to exert coercion because more votes are prevented from going to the alternative candidate as a result of the scandal when ϕ_m and μ_m are large.

Instead, it is less likely that after the media scandal a parapolitician obtains more votes in municipality m if ψ_m is high, as state capacity increases the marginal cost of coercion, reducing incentives to compensate for the media scandal with increases in coercion. Also, when the population of municipality m is strongly biased ideologically towards A (when σ_m is large), the response of coercion to the media scandal will also be weaker because the population is inclined to vote, even freely, for the parapolitician.

Since the media scandal need not reduce the vote share for paramilitaries in any given municipality m , total electoral support for the parapolitician following a media scandal does not necessarily fall either. This is the second key result in Proposition 2: media exposure of parapoliticians does not make them more electorally accountable when they can “hide in the shade”. Of course, this does not mean that parapolitician wellbeing is unaffected by the media scandal. It is straightforward to show that following a national media scandal, politician A is worse off, even if he can sustain his electoral success⁹.

⁹Direct differentiation of politician's A equilibrium welfare $\Pi^A - C(\{\lambda_m^*\})$ shows that

$$\frac{d\Pi^A}{d\mu} - \frac{dC(\{\lambda_m^*\})}{d\mu} = - \sum_m l_m (1 - \lambda_m^*) \phi_m < 0,$$

which is simply an application of the Envelope's Theorem.

Needless to say, a number of simplifying assumptions drive our results. Indeed, more than proposing a very general model of electoral competition under coercion and media exposure, we attempted to present the simplest model to capture the idea that an endogenous response in coercion may at least partly compensate the electoral impact of a media scandal, especially in areas with a weak institutional environment. In Appendix A, we consider one particularly important extension. In the baseline model, an increase in media exposure of the parapolitician only affects his popularity. However, one could conceivably argue that more media exposure also makes coercion more costly. In the Appendix, we incorporate this additional ingredient and show that, as long as coercion costs do not increase linearly with the media scandal, the results from the simpler baseline model are unaffected.

2.3 Discussion and anecdotal evidence

This simple framework has a number of important implications. First of all, equation (25) shows that, other things being equal, parapoliticians (A) have an edge over non-parapoliticians (B) because they exert coercion. Second, equation (2) shows that, as a result of coercion, paramilitary-controlled areas tend to receive less than full attention in terms of policy, and that this attention decreases with increases in coercion. The intuition is simple: coerced voters do not respond to policy, so there is no electoral incentive to cater to their demands.

While lack of good quality data on municipal-specific policies makes it difficult to find evidence supporting the second prediction, there is a large bulk of anecdotal evidence supporting the fact that parapoliticians have an advantage over non parapoliticians. One telling case study is that of politician Alvaro Araujo (see Acevedo (2010)). Coming from a family with a longstanding presence in the political arena of the department of Cesar, Araujo was elected to the House of Representatives for two consecutive legislative periods, in 1994 and 1998. Araujo’s votes were then concentrated in northern Cesar, the political stronghold of his family. In 2002 Araujo was elected Senator with a geographical vote pattern substantially different than the one he had during his previous experience in the House: most of his votes were concentrated in southern Cesar, where his family traditionally had very little electoral support. Instead, this region was controlled by the Northern Block of the AUC, headed by the paramilitary warlord “Jorge 40”.¹⁰ In 2007 Araujo was found guilty of having ties with Jorge 40. It was then proved that the men of Jorge 40 ensured the election of Araujo (as well as that of other politicians under their sponsorship) through a mix of fraudulent practices and the intimidation of voters in their region of influence. Particularly telling is the kidnap of Araujo’s political competitor Victor Ochoa. Moreover, another rival political family from the region, the Gneccos, was effectively banned from politics through threats and intimidation.

A second case study of interest is that of politician Juan Carlos Martinez, from the department

¹⁰Extradited to the US in 2008 facing drug trafficking charges, Jorge 40 is more famous in Colombia for his participation in massacres, selective homicides and the forced displacement of civilians.

of Valle (Cruz, 2010). Martinez’ links with the paramilitaries were revealed by paramilitary leader “HH” in 2008. According to HH’s, Martinez was elected Senator in 2002 largely with the support of the Calima block of the AUC. However, after the demobilization of this block in 2004 in the context of the AUC negotiation with the government, Martinez lost support in the areas controlled by the paramilitary in the south-western part of the country, notably Tumaco, Zarzal, Candelaria and Tulua. In an effort to offset this loss, Martinez looked for the support of alias “Rasguño”, the chief of the “Norte del Valle” drug cartel (Laverde, 2012). Thus, Martinez’ geographical distribution of votes changed substantially from 2002 to 2006 and, in spite of the demobilization of his original illegal sponsors, Martinez was re-elected in the Senate with 3,354 more votes relative to 2002. Martinez’ 2006 votes were obtained in areas controlled by the drug cartel.

The more important implications of the model, however, concern the impact of the media scandal and the endogenous response of parapoliticians (propositions 1 and 2). Despite the existence of coercion, a natural intuition is that a media scandal revealing ties of paramilitaries with politicians ought to bring greater political accountability. The framework above shows that this intuition is not necessarily true. In particular, our results highlight that there are three ways in which a media scandal may have negative unintended consequences. First, it increases coercion. Second, as a result, politicians pay less attention to voters’ needs. Third, if this endogenous response of coercion is sufficiently strong, the media scandal may not even reduce the overall vote share for politicians that are shown to have links with paramilitaries, an arguably desirable objective of the media revelation, and what Jefferson and Brandeis had in mind.

This is also supported by a large body of journalistic accounts. Take again, for instance, the case of politician Alvaro Araujo. Even if Araujo’s 2002 Senate campaign had the support of Jorge 40 and the Northern Block of the AUC, this was not made public until well after Araujo was elected. The first accusations linking Araujo and Jorge 40 came from journalist Claudia López in 2005 (Semana, 2006b). As a result, in the 2006 elections Araujo lost most of his electoral support in places with a stronger institutional background. For example, compared to the 2002 elections, Araujo lost 12,000 votes in capital city Bogotá. However, he reached again for the support of Jorge 40, who ordered his men to get him votes in the department of Atlántico in addition to Cesar. This strategy was successful and Araujo went from 146 votes in 2002 in Atlántico’s capital (Barranquilla) to 6,752 in 2006. Overall, having been exposed as parapolitician before the 2006 elections, Araujo only lost in total 2,848 votes relative to 2002. He was elected Senator for the second time.¹¹

These simple results highlight the complementarity between the different dimensions of insti-

¹¹Other incumbent Senate candidates had similar trajectories. For instance, Dieb Maloof, Vicente Blel and Habib Merheg all were accused by the media of their ties with well known paramilitaries (respectively Jorge 40, Diego Vecino and Salvatore Mancuso, and alias “Rasguño” and “Macaco”) before the 2006 elections. As a result, even if the three politicians had belonged to the Liberal Party in the past, they were denied the opportunity to use the name of that party in the elections, or that of any other of the main political parties for that matter. The politicians founded their own party, “Colombia Viva” and were elected to the Senate with votes concentrated in paramilitary-dominated areas of the departments of Magdalena (Maloof), Bolivar (Blel) and Risaralda (Merheg).

tutions in a democracy. Having a free, active media may not be sufficient and may even bring undesirable unintended consequences if other dimensions of institutional quality, like the existence of free and fair elections, are not guaranteed. Indeed, the framework above suggests that these negative consequences of the media scandal would be reduced with sufficient state capacity (sufficiently high ψ). With this framework in mind, we next turn to our data and empirical strategy, and complement the anecdotal discussion of this subsection with systematic quantitative evidence.

3 Data

A key challenge when testing the implications of the parapolitics scandal is that, because of its illegal nature and its reliance on threats and intimidation, it is very hard to observe coercion at the local level and measure it systematically. We overcome this constraint by focusing on the testable electoral consequences of the scandal derived in our theoretical model. Our key dependent variable is each senate candidate’s municipal vote share in each election year from 2002 to 2010. The source of electoral results’ data is *Registraduría Nacional del Estado Civil*, the national electoral authority.¹² On the other hand, to capture the parapolitics scandal we constructed an original dataset of news stories connecting national politicians with paramilitaries using the online archive of Colombia’s largest newspaper with nation-wide circulation and coverage, *El Tiempo*. We now describe the coding process and other key variables in the analysis, with further details discussed in Appendix B.

3.1 Media exposure of parapoliticians

Our media exposure dataset is event-based and relies on information on every news story (and op-ed) published from January 1st 1997 to August 1st 2011 in *El Tiempo*. We searched manually (relying on human coders for greater accuracy) for stories in every section of the newspaper about each of the incumbent congressmen in this 15-year period. For every story we record: the date, the newspaper section in which it appears, and whether it links the politician with parapolitics activity.

We focus on senators, elected in Colombia since the 1991 Constitutional reform on the basis of a single nationwide district. This reform, as well as subsequent changes in the electoral system in Colombia, have sought to create incentives for senators to appeal to a broader electorate (Roland & Zapata, 2004). While senators may still be elected successfully by obtaining a large share of their votes from a few municipalities or regions (and they often do), they capture more national attention and votes than candidates to the House. House members are elected on the basis of regional districts (each of Colombia’s 32 Departments, the equivalent of the US State) and a few special districts (such as the Capital District of Bogotá). Hence, they receive much less national

¹²From the same source, we compute turnout in each municipality and election year as the ratio of total votes to the municipal voting age population.

attention and, crucial for our empirical strategy, coverage in national newspapers than national senators. To have a set of more comparable candidates, we focus on senators with at least some previous experience as members of the Senate or the House of Representatives.

Based on our news search, we construct three key measures of the connectedness of each politician with the parapolitics scandal. First, we look at the extensive margin by coding a dummy variable that equals 1 if there are any news stories connecting the candidate to paramilitaries at any point in time. Second, because some politicians are more deeply involved in the scandal than others, we exploit the intensive margin by coding the percentage of total news stories of each candidate that relate him with the scandal.¹³ Finally, it is key to distinguish parapoliticians exposed to the scandal before an election from those exposed afterwards. Thus, we also code a dummy variable that equals 1 if the candidate was linked to paramilitaries before each of the elections in our sample period (2002, 2006, and 2010).

Our method to identify a parapolitician has a number of advantages relative to those that have been used in the literature, particularly given our interest on the effects of media exposure. For instance, since several parapoliticians belong to ‘third parties’ (that is, parties other than the Liberals, Conservatives, and the Socialists), some papers have relied on the rise of these non-traditional third parties in certain areas of the country as a proxy of the strength of parapoliticians (e.g. López (2007), Valencia (2007), and Acemoglu et al. (2013)). Nevertheless, there is significant measurement error in this proxy for being a parapolitician. On the one hand, some politicians in the traditional parties have been linked to paramilitaries, both by the press and by judicial institutions. On the other hand, some of the third parties, far from being dominated by politicians connected to paramilitaries, are in fact independent parties seeking alternatives to the traditional parties (perhaps the most notable example is the ‘Partido Verde’, which was particularly important player in the 2010 presidential campaign and also elected 5 senators in 2012). So, although as demonstrated by Acemoglu et al. (2013) on average the vote share for third parties in legislative elections is indeed correlated with paramilitary presence, there is considerable noise in this relation. Our media-based individual-level measure is much more precise and, of course, better suited for our specific purpose of verifying the implications of media exposure.

The coding of stories from *El Tiempo* provides a very precise measure of the national media scandal (μ in our model). While with differing degree of emphasis or even tone, we expect the stories that we capture in our news search to have been reproduced at the local level (not just in newspapers, but also in both national and local radio stations and TV broadcasters). Direct evidence that the news in *El Tiempo* proxies not just what this newspaper was presenting to the audience but what other media outlets were uncovering is presented in Section 5.3.3, where we compare *El Tiempo* with another national newspaper, *El Espectador*.¹⁴

¹³We normalize this variable by total candidate stories as some candidates are much more popular than others, and get more press coverage.

¹⁴Newspaper penetration, and the extent to which stories are reproduced (or not) in the local press, radio, and TV

3.2 Shade and sunlight

To test the implications of our theory, we focus on municipal variation on factors that make coercion more ('sunlight') or less ('shade') costly. We rely on paramilitary presence at the municipal level as the main 'shade' variable since parapolitics is precisely about politicians outsourcing vote manipulation to illegal paramilitary groups. For 'sunlight', we rely on measures of state capacity that increase the cost of exerting coercion.

To measure paramilitary presence, we use the index from Acemoglu et al. (2013). These authors rely on paramilitary attacks from the research center CEDE at Universidad de Los Andes in Bogotá, as well as data on internally displaced people (IDP), forced by paramilitaries, from Colombia's Victims Unit.¹⁵ Because the time series variation of both paramilitary attacks and IDP is quite noisy, Acemoglu et al. (2013) focus on municipal-specific cumulative measures over the period 1997-2005. We borrow two measures of paramilitary presence: an index that extracts the principal component of the attacks and IDP measures, and a dummy that takes the value of 1 if the cumulative paramilitary attacks of a municipality are in the top 30% of the distribution.¹⁶ All our results are robust to using other measures of paramilitary presence, that rely on a completely independent data source, described in Dube and Vargas (2013).

To measure state capacity, we compute two proxies at the municipality level using data from *Consejo Superior de la Judicatura*. First, we use principal components analysis to compute an index that summarizes data on the number of judges, attorneys and general prosecutors stationed in each municipality. In this sense our measure captures the presence of law enforcing institutions at the municipal level. We apply the index over the standardized variables, and keep the first principal component, which alone explains over 96% of the variance. Second, we code a dummy variable that equals one if the number of judges, attorneys and prosecutors per capita is in the top

could provide additional variation in the intensity of the scandal (l_m in our model). Unfortunately, however, data limitations precludes taking advantage of this additional source of variation. Even circulation figures for *El Tiempo* and other national newspapers are not regularly published or verified in Colombia. Unlike in the US where the industry complies with the Audit Bureau of Circulation producing a reliable record, in Colombia the industry converged to using surveys to have independent measures of penetration. These surveys only focus on a few cities. While useful for most newspapers' purposes (e.g. for selling advertisement), this information is insufficient to run regressions at the municipality level. Moreover, even though some figures (of disputed quality) are available for national newspapers, circulation of local newspapers which are important in municipalities outside the capital, Bogotá, is even more incomplete. Finally, local newspapers also have very limited electronic archives to build a measure of exposure using the regional press as we do with the national press using *El Tiempo's* electronic archive.

¹⁵CEDE collects data from the Observatory of Human Rights of the Vicepresidency and the National Department of Planning and aggregates variables in several categories by armed actor (paramilitaries and guerrillas) and type of action. Attacks are coded as the sum of explosive terrorist acts, incendiary terrorist acts, assaults to private property, attacks on civil organizations, political assassination attempts, road blockades, ambushes of civilians, incursion into villages, overland piracy, illegal checkpoints, murder of civilians, murder of politicians, massacres, kidnappings of members of the armed forces and kidnappings civilians. Data on displaced population are collected from the Registro Único de Víctimas (unique register of victims) and specify the municipality where the displacement originated, the year of the displacement and the identity of the perpetrating group.

¹⁶The reason of the first strategy is that both attacks and IDP numbers are noisy, thus their common component may contain more information.

30% of the distribution.

3.3 Other data

We use measures of guerrilla activity for a placebo exercise, since the parapolitics scandal revealed the links of politicians with illegal armed right-wing paramilitaries, not with left-wing guerrillas. This implies that if the vote share of parapoliticians is higher than that of non-parapoliticians in places with guerrilla activity (but not paramilitary activity), then it may be the case that parapoliticians' votes are concentrated where the armed conflict is more intense, rather than where paramilitary activity (and hence the likelihood of electoral coercion by paramilitaries) is large.

We consider two variables which aim to identify places of high guerrilla presence. Our first measure comes from López (2010): a dummy variable that equals 1 if there is evidence (investigations and judicial processes, ongoing or finalized) of political linkages between local politicians of a municipality (councilmen or mayor) and guerrillas. Thus this is a measure of the capture of local institutions and decision-making by guerrilla groups. The information for this variable comes from a rich set of regional studies conducted by the NGO *Corporación Nuevo Arco Iris*. We also construct a guerrilla presence index equivalent to the one described for paramilitaries and code a dummy variable that equals 1 if the index is in the top 30th percentile of the distribution. Since guerrilla activity is correlated with paramilitary activity, in these placebo tests we control for paramilitary presence as described above.

We include a set of individual-level controls for the politicians in our dataset. A key control is the total number of stories about the politician from our own press search, which may capture both the overall popularity of the candidate and his national exposure in the press. From *Congreso Visible*, an NGO that collects information on the Colombian Congress from multiple sources, we coded each politicians' gender, profession, whether the politician is a former member of the House or the Senate, his years in politics, the number of periods in office, and a dummy variable that equals one if the senator did not complete his last term (for reasons such as having been discharged from Congress for misbehavior).

4 Empirical Strategy

4.1 Main specification

Our estimation strategy exploits variation at the municipality \times candidate level in each one of the three electoral years when the practice of colluding with illegal armed paramilitaries took place: 2002, 2006 and 2010. This allows us to compare the municipal vote share of candidates exposed and not exposed by the media scandal (as well as the vote share of parapoliticians exposed before the elections relative to that of parapoliticians exposed after elections), and test whether this

difference varies according to the relative institutional weakness/strength (‘shade’ or ‘sunlight’) of the municipality. We estimate:

$$\pi_{cmt} = \alpha + \omega_m + \delta_t + \beta_1 \text{Exposed}_c + \beta_2 (\text{Exposed}_c \times \text{Shade}_m) + \beta_3 (\text{Exposed}_c \times \text{Sunlight}_m) + \theta' \mathbf{X}_c + \varepsilon_{cmt}. \quad (8)$$

where π_{cmt} is the vote share of candidate c in municipality m and election year t . Exposed_c is one of our measures for whether candidate c is a parapolitician, and Shade_m (Sunlight_m) is one dimension of ‘shade’ (‘sunlight’) at the municipal level, which influences the cost of exerting coercion to obtain votes in municipality m . Thus, β_1 captures the differential electoral success of parapoliticians relative to ‘clean’ candidates, which according to our model should be positive (because of candidate’s A coercion capacity). In turn, β_2 (β_3) captures the extent to which the parapolitician’s electoral advantage is either boosted (hindered) by the relative institutional weakness (strength) of the municipality. Our model predicts that parapoliticians should have more votes in areas with high paramilitary presence and weak state capacity (small ψ_j in the model). Hence, we expect β_2 to be positive and β_3 to be negative.

We include a full set of municipality (ω_m) and time (δ_t) fixed effects to allow for municipal-specific, time-invariant heterogeneity and for aggregate heterogeneity across election years.¹⁷ Also importantly, since being a parapolitician may be correlated with other candidate characteristics, we include a large set observable candidate-level controls in the vector \mathbf{X}_c . Finally, ε_{cmt} is the error term.¹⁸

Equation (8) is a useful description of the patterns of electoral support for parapoliticians and non-parapoliticians, unveiling correlations that can be interpreted under the lens of our theory.¹⁹ However, the more interesting implications of our model concern the impact of the media scandal vote distribution. In particular, we expect that following exposure by the media, parapoliticians exacerbate their efforts to obtain votes in coerced areas, especially in the context of weak institu-

¹⁷From 2003 to 2006 large numbers of paramilitaries (allegedly around 35,000) demobilized collectively in the context of negotiations between the paramilitary umbrella organization AUC and the government of president Uribe. However, paramilitaries hardly lost their capacity of exerting coercion in parts of the Colombian territory, both because some paramilitary blocks refrained from participating in the process, and because several of the blocks that did quickly resorted to arms and formed what is now known as “neo-paramilitary” criminal bands. Indeed, Figure 2 shows that, in spite of a downward trend starting in 2003 (when the demobilization process started), attacks by paramilitary groups peaked in 2005, just before the (March) 2006 Congress elections which drive most of our empirical results.

¹⁸We report standard errors clustered at the municipality level and two-way clustered at the municipality and politician level. The first type of clustering allows for serial correlation over time within municipalities. The second corrects for potential serial correlation within municipalities and politicians (some of which participate in more than one election). The two-way clustering produces larger standard errors throughout, but the significance of the substantive effects (coefficients β_1 , β_2 and β_3 in 8) is largely maintained.

¹⁹A potential concern is that some other municipality characteristics, that correlate with either ‘shade’ or ‘sunlight’, explain the differential vote share of exposed parapoliticians captured by β_2 and β_3 . We run specifications similar to (8) (and to (9) below) that also control for some municipal characteristics such as basic geographic controls and preexisting socioeconomic outcomes, interacted with the exposure dummy. The results (available upon request) are robust to those reported for the simpler specifications in (8) and (9).

tions. Hence, we also run the following regression, similar to equation (8) but only for the set of candidates that we identify as parapoliticians:

$$\begin{aligned}\pi_{pmt} = & \alpha + \omega_m + \delta_t + \tilde{\beta}_1 \text{Exposed before elections}_p + \tilde{\beta}_2 (\text{Exposed before elections}_p \times \text{Shade}_m) \\ & + \tilde{\beta}_3 (\text{Exposed before elections}_p \times \text{Sunlight}_m) + \theta' \mathbf{X}_c + \xi_{pmt}\end{aligned}\quad (9)$$

where $p \in c$ denotes the subset of senators in our sample that were linked by the media with paramilitaries at any point in time. Our key independent variable, $\text{Exposed before election}_p$, is a dummy variable that equals 1 if parapolitician p was linked to the scandal before the election. In this specification, the theoretical prediction of the effect of exposure prior to an election on vote share is ambiguous, and depends on whether parapoliticians are able to compensate via coercion their reduced popularity. The interesting insight, however, is that $\tilde{\beta}_1$ need not be negative, since the endogenous response of coercion may avoid an electoral cost of the media scandal. Again, measuring Shade_m with paramilitary presence of the areas where candidate p obtained votes, we expect $\tilde{\beta}_2 > 0$; measuring Sunlight_m with local state capacity, we expect $\tilde{\beta}_3 < 0$.

Since this regression only includes parapoliticians, it allows us to check the effect of the media scandal, excluding the overall differences in the patterns of electoral support between parapoliticians and non-parapoliticians that (8) explores. However, even in this case a threat to identification is that, within the group of parapoliticians, those exposed before elected may differ systematically to those exposed after elected. Unobserved ability or other candidates characteristics could therefore bias our results. To assuage this concern, we show that the set of parapoliticians exposed before elections have similar observable characteristics as those exposed once elected.²⁰

4.2 Alternative specification

Given that our theoretical framework predicts that parapoliticians should have a different distribution of votes relative to non-parapoliticians (and that such distribution is likely to be affected by the media scandal), an alternative approach is computing candidate-specific aggregates of the extent to which they receive support in areas characterized by sunlight or shade, and correlate them with our individual measures of exposure to the scandal. A potential advantage of this strategy is that candidate performance averages might remove part of the idiosyncratic noise in municipal level outcomes. Specifically, for each candidate c and election period t , we compute:

²⁰An alternative would be to follow a regression discontinuity approach, thus comparing only the set of candidates exposed by the media “shortly” before and “shortly after” election day, controlling flexibly for distance (days, weeks, or months) to elections. Unfortunately, however, we do not have a large enough sample of politicians to follow this regression discontinuity approach. The number of parapoliticians competing in our elections is ‘only’ 77, corresponding to 119 observations through our three election years.

$$\begin{aligned}\overline{\text{Shade}}_{ct} &= \sum_m \pi_{mt}^c(\text{Shade}_m) \\ \overline{\text{Sunlight}}_{ct} &= \sum_m \pi_{mt}^c(\text{Sunlight}_m)\end{aligned}\tag{10}$$

where $\text{Sunlight}/\text{Shade}_m$ is a variable capturing municipal ‘shade’ or ‘sunlight’ characteristics, that according to our model should affect the distribution of votes for parapoliticians. We weight this municipality characteristics by π_{mt}^c , the vote share of candidate c (out of his own total votes) in municipality m at time t . Hence, these weighted averages summarize the relevant characteristics of the municipalities where candidate c got his votes from at time t . We then estimate:

$$\overline{\text{Shade}}_{ct} = \alpha + \delta_t + \hat{\beta}_1 \text{Exposed}_c + \theta' \mathbf{X}_c + \varepsilon_{ct}.\tag{11}$$

where the right hand side elements are interpreted as in (8), and similarly for $\overline{\text{Sunlight}}_{ct}$.²¹

As demonstrated in Appendix C, with this alternative empirical specification we can also interpret β_1 in regression model (11) (and its equivalent within parapoliticians) using our theoretical model.

5 Results

Tables 1A and 1B present descriptive statistics of all our variables (at the municipality and at the candidate level, respectively). The latter includes descriptive statistics of candidate characteristics for parapoliticians (involved in the scandal at any point in time) and non-parapoliticians. From this comparison it is apparent that parapoliticians have a 23 percent higher vote share than non-parapoliticians. This difference, which is statistically significant, is suggestive and we will explore it more systematically. Also, individual-level candidate controls are largely balanced. This suggests that selection on observables is unlikely to drive the results obtained when comparing parapoliticians exposed before the elections with those exposed afterwards.²²

5.1 Coercion and the electoral performance of parapoliticians

Table 2 reports the results from estimating the specification in equation (8). Each one of the measures of exposure to the parapolitics scandal (and its interaction with the shade and sunshine municipality characteristics, described in the previous section) is reported in one panel. Panel A

²¹As with municipal-candidate-level approach (see (9)), we also estimate a version of (11) within parapoliticians p , using as exposure our measure of exposure before elections.

²²Table C-1 in Appendix C shows descriptive statistics for the variables used in the alternative candidate-level regressions.

focuses on the (extensive margin) exposure dummy, that takes the value 1 for politicians exposed to the scandal at any point during the sample period. Panel B reports the (intensive margin) measure of share of news about the scandal to the total news about a politician. In columns 1 and 2 each measure of exposure is interacted with the two proxies of paramilitary presence described in section 3. In columns 3 and 4 they are interacted with the proxies of state capacity. Columns 5 and 6 estimate the most demanding specification, in which both interactions are included.²³

Controlling throughout for municipality and year fixed effects, as well as for candidate-specific observable characteristics, we find that the non-interacted exposure measure is positive and significant.²⁴ This is true for any of the three proxies of exposure to the scandal, and it is consistent with the theoretical model’s prediction that, because of their alliance with illegal paramilitaries, all else equal parapoliticians have an advantage over non-parapoliticians in terms of electoral performance.

Depending on the specification, in the extensive margin (Panel A), parapoliticians can have a vote share margin from 0.06 to 0.12 percentage points higher than non-parapoliticians. This magnitude is large: the vote share edge of parapoliticians is between 13 and 26 percent of the mean vote share and between 2 and 4 percent of its standard deviation (see Table 1A).

Also, in the extensive margin parapoliticians have an even larger vote share margin (from 0.04 to 0.15 percentage points larger depending on the specification) in places with paramilitary presence. This is significantly different from zero independently of the measure of exposure and the proxy of paramilitary presence, and regardless of whether the standard errors are clustered at the municipality or at the municipality/politician level (see footnote 18). However, but still consistent with the theoretical model, in places with higher state capacity, parapoliticians are not able to exert so much coercion and thus their electoral advantage relative to non-parapoliticians is curtailed, although not enough so as to completely eliminate the vote share margin of parapoliticians.

Table C-2 in Appendix C reports the equivalent results using the alternative candidate-level specification described in section 4. The results are qualitatively equivalent: parapoliticians have a positive vote share margin relative to non-parapoliticians²⁵, a greater concentration of votes in ‘shade’ municipalities (where the cost of coercion is lower) and their votes are less concentrated in ‘sunlight’ municipalities (where the relatively high state capacity makes coercion more costly).

²³Column 5 uses the continuous indices of paramilitary presence and state capacity and column 6 the dummies of high paramilitary presence and high state capacity.

²⁴Note that the non-interacted municipal characteristics of ‘shade’ or ‘sunlight’ is omitted because it is collinear with the municipality fixed effect.

²⁵This is particularly important because in the municipality-level specifications (models (8) and (9)), the coefficient for exposure to the scandal (β_1) does not exactly measure overall parapolitician advantage. Indeed, β_1 is the unweighted average vote share advantage of exposed candidates at the municipality level, but each municipality contributes differently to the *total* vote share depending on the municipality’s total votes. In the candidate-level regressions discussed in section 4.2, the dependent variables are candidate-level aggregates. In Column 1 of Table C-2, in particular, we simply look at performance: each politician’s total vote share. The results confirm the theoretical predictions that parapoliticians have an overall electoral advantage.

5.2 The media scandal and the electoral performance of parapoliticians

We now estimate regression model (9) only for the sample of parapoliticians. Our key independent variable of interest is exposure before the election: the dummy variable that equals 1 if the candidate was linked to the parapolitics scandal before the election at time t . As noted, since this regression includes only the set of parapoliticians, it allows us to check the effect of the media scandal, excluding the overall differences in the patterns of electoral support between parapoliticians and non-parapoliticians that might contaminate the results of Table 2.

For the coefficient on exposure to measure the effect of the media scandal, we need the set of parapoliticians exposed before elections to be otherwise comparable to the set of parapoliticians exposed afterwards. The inclusion of candidate-level controls (in addition to the municipality and election-year fixed effects) helps dealing with the potential differences between parapoliticians who manage to stay under the radar before the elections and those who do not. Moreover, we can examine whether there are any observable significant differences between these two sets of parapoliticians. This is done in Table 3, which reports estimated coefficients from a linear probability regression model of the dummy of exposure before elections on our set of observable candidate controls as well as election-year fixed effects. This specifications allows us to explore whether observable individual characteristics of parapoliticians predict exposure to the scandal before elections, within each election year.

It is reassuring that most of the observable candidate controls are balanced across parapoliticians exposed before and after elections. Only the average periods in congress (at the 10% level) and the total number of news stories in *El Tiempo* (at the 5% level) are significant (Column 1). However, the magnitudes of the differences are relatively small, including for periods in congress that comes out significant (about 0.4 is the estimated difference, for an average number of periods of 2.1 in the sample). Moreover, if we examine the types of news stories and explore the number of non-paramilitary related news stories only, the estimated coefficient is no longer significant (Column 2). This suggests that the difference in total news stories corresponds to the parapolitician exposure itself. That is, instead of differences in popularity or newsworthiness between politicians exposed before or after the elections, the difference in the number of total news stories simply reflects that exposed parapoliticians mechanically have more stories, as their links with paramilitaries are being reported.

Overall, therefore, we take Table 3 as giving evidence that our results are unlikely to be driven by selection on observables between parapoliticians exposed before and after elections. In any case, and to have more precise estimates, we control for all observable characteristics in the regressions that follow.²⁶

Table 4 shows the results obtained when estimating specification (9), which compares para-

²⁶The magnitude of estimated coefficients is not affected by the inclusion of these controls.

politicians exposed before and after elections. Recall that our framework implies that the effect of exposure can be positive, negative or null. This is because the media scandal shifts the preferences of voters against parapoliticians, but the endogenous coercive response offsets the electoral cost of the scandal. This is precisely what we find: regardless of the specification parapoliticians exposed before elections do not have an electoral cost (or advantage) relative to those exposed afterwards. While negative in sign, the estimated coefficient is small and not significant.²⁷

We also find evidence that exposure before elections increases the vote share margin of parapoliticians in areas with paramilitary presence (significant only for the continuous measure of paramilitary presence, in columns 1 and 5). This result is consistent with the mechanisms of our model: the media scandal prompted parapoliticians to double-down on coercion in places where this was less costly, and by doing this they compensated for the overall fall in popularity among free voters.

In contrast, increases in municipal state capacity do not appear to ensure that parapoliticians exposed before elections are penalized by voters (columns 3, 4 and 6). The interaction coefficients are positive, small, and not significant. While this is at odds with the model, when we focus on the alternative candidate-level specification of Appendix C, we do find that parapoliticians exposed before elections have a vote pattern much less concentrated in high state capacity areas, relative to parapoliticians exposed after elections (see Table C-3). This might be because municipal heterogeneity inherent to the candidate-municipality specification introduces noise that is removed when aggregating at the candidate level. However, to be on the conservative side, we use regression models (8) and (9) as our baseline and interpret the evidence on the role of state presence for parapoliticians exposed before elections with caution.

5.3 Further robustness

5.3.1 Parapoliticians and unusually high turnout

Table 5 presents an additional piece of evidence that supports the purported underlying mechanism that explains the vote share edge of parapoliticians. We estimate again regression models (8) and (9) but interact each one of the exposure measures with a dummy variable that equals one if a municipality had unusually high turnout in legislative elections. Specifically, high turnout is defined as equal to one if the municipal turnout (the ratio of total votes to voting age population) is in the 30% upper decile of the turnout distribution.

Columns 1 and 2 use as exposure proxies the extensive margin and the intensive margin measures respectively, thus comparing parapoliticians to non-parapoliticians (specification (8)). Column 3

²⁷Focusing on the candidate-level results, the equivalent Table C-3 of Appendix C also shows (column 1) that the vote shares of parapoliticians exposed before and after elections are similar (the coefficient for exposure is positive but not significant). As argued in footnote 25, this specification captures overall electoral success more precisely than the municipal level regressions which give equal weight to large and small municipalities.

focuses on the dummy for exposure before elections, thus comparing parapoliticians exposed before and after the elections (specification (9)).

As previously, we find that parapoliticians get a higher vote share than non parapoliticians, irrespective of the measure of exposure (Columns 1 and 2). Moreover, we also find that the positive vote share margin of parapoliticians is larger in municipalities that experience unusually high turnout during legislative elections.²⁸

These results lend further support to the idea that paramilitary coercion is the underlying mechanism explaining our results. In particular, while there are various (fraudulent) methods that politicians in general could use to get electoral support, and to offset the negative impact of a media scandal, anecdotal and judicial evidence in the case of parapoliticians in Colombia points to methods of coercion that would tend to inflate vote totals.

Some of these methods were already mentioned in our discussion of the anecdotal evidence. Acemoglu et al. (2013) also offer a useful summary of the parapoliticians fraudulent practices. As they point out, “A salient strategy seems to have been to terrorize people into voting for specific candidates. In the municipality of San Onofre in the coastal department of Sucre, for example, this was arranged by the paramilitary leader ‘Cadena’: for the elections of 2002, the trucks sent by ‘Cadena’ went through neighborhoods and rural areas of San Onofre picking people up. According to some people in this municipality in Sucre, thousands of peasants were taken to the corregimiento ‘Plan Parejo’ so they could see the candidate for whom they had to vote for in the legislative elections.” Admittedly, terror was also used to keep people away from the polls, but this was done so “that ballot stuffing and other forms of manipulation of vote totals could occur.” Moreover, “another strategy, where coercion also played an important role, involved collecting people’s cédulas (national identity cards which a person must produce to vote) from their houses, using them to collect the ballots (the ‘tarjetón’) and filling them in for people.” All of these strategies increase the likelihood of unusually high turnout in places where parapoliticians could exert coercion, and Table 5 indeed suggests that parapoliticians got votes from places with unusually high turnout areas.

We do not claim, however, that every form of parapolitician fraud involved armed coercion, nor that it implied increased vote totals. For instance, in judicial investigations it was revealed that the Presidential intelligence service (Administrative Security Department, DAS), now closed after these and other scandals, designed a computer program to create counterfeit ballots to include fake votes of the people who did not vote, and to replace those of voters who did turn out to vote for candidates allied with paramilitaries. However, since most strategies did imply increased vote totals, and especially in the places where coercion is more likely, we expect the positive association

²⁸While parapoliticians exposed before elections also get a higher vote share in high turnout municipalities than those exposed after elections, in this case the difference is not statistically different from zero (Column 3). However, using the alternative candidate-level specification, Table C-4 of Appendix C does suggest a significant effect when comparing parapoliticians exposed before and after elections.

that is apparent in Table 5.

5.3.2 High guerrilla activity areas: a placebo

We posit and formalize the idea that paramilitary presence facilitates the coercion of voters into supporting their preferred candidates. Further, Tables 2 and 4 (as well as Appendix C’s Tables C-2 and C-3) support this prediction. However, an alternative explanation of the empirical patterns is that parapoliticians’ votes are just concentrated where the Colombian armed conflict was strong. For example, since guerrilla activity is correlated with paramilitary activity, we may be spuriously finding a correlation that really stems from guerrilla, not paramilitary, presence.

In this section, as a placebo, we re-estimate regression models (8) and (9) but interact each one of the exposure measures with the two different proxies of guerrilla activity described in section 3. The idea behind this exercise is to test whether parapoliticians vote share margin is higher in places of guerrilla dominance, controlling for paramilitary presence. If that was the case, our hypothesized mechanism would be proven false.

Table 6 presents the results for all politicians (Panels A and B) and within parapoliticians (Panel C). We find further support for the model’s predictions, as parapoliticians do not get a larger vote share margin relative to non parapoliticians in areas with high guerrilla presence. Rather, the opposite is true: guerrilla activity hinders the coercive capacity of politicians that collude with illegal paramilitaries (which are the guerrilla’s enemy). Importantly, in this specification the coefficients of the interactions of exposure to the scandal and paramilitary presence remain significant when comparing parapoliticians to non-parapoliticians. We obtain similar results when comparing parapoliticians exposed before elections to those exposed after. Finally, Table C-5 of Appendix C confirms these results using a different specification.

5.3.3 Checking media bias

We used the electronic archive of *El Tiempo*, the largest newspaper with nation-wide coverage in Colombia, to code the extent to which every incumbent Senate candidate was allegedly involved with the parapolitics scandal. We rely on this media outlet for substantive and practical reasons. First, it is the most important newspaper of national circulation. Second, it is the only one with an electronic archive going back beyond the period in which the media scandal broke out. In particular, *El Tiempo*’s electronic archive currently goes back to 1991.

Focusing on this one paper is, however, a potential source of concern. Indeed, *El Tiempo* has traditionally been viewed as a pro-government, center-right newspaper. Moreover, when 55% of the paper’s shares were bought by Spain’s media conglomerate *Grupo Planeta* in August 2007, the perception that *El Tiempo* further leaned to the right was not uncommon.

This is potentially problematic if *El Tiempo*’s editorial position affected the objectivity of its

news coverage during the period of analysis. For instance, *El Tiempo* supported Uribe’s reelection in 2006, at the time the parapolitics scandal was gaining momentum (see Figure 1) and most of the incumbents then exposed as parapoliticians belonged to the president’s coalition. It may be argued that, by relying on *El Tiempo* to code the intensity of the involvement of incumbent senate candidates with the scandal we are underestimating the scope of the parapolitics phenomenon. This in turn may be a source of bias for our results.

To address this possibility we coded the news coverage of all incumbent Senate candidates and their alleged involvement with parapolitics in *El Espectador*, which is the second largest nation-wide newspaper in Colombia. *El Espectador* has a reputation of independence, and a track record of exposing political (and other kind of) misdeeds. For instance, *El Espectador* denounced the drug mafias in the 1980s and their links with politics like no other medium. In 1986, the paper’s lifetime director, Guillermo Cano, was murdered by Pablo Escobar in retaliation. Also, in 2007, marking a fundamental dissociation with *El Tiempo*’s editorial stance, *El Espectador*’s director criticized president Uribe (Semana, 2007).

Using the *El Espectador* data, we can contrast the *El Tiempo* coverage on parapolitics-related news stories for all the politicians in our sample. One caveat is that the electronic archive of *El Espectador* is not available before October 2007. Hence for the comparison we rely on the overlapping period of available information from the two papers, spanning between October 2007 and August 2011. Candidates exposed only by *El Tiempo* before this period, are dropped from our regressions.

For each paper we rank our sample of politicians according to the ratio of parapolitics-related stories to the total number of stories on each individual. We use this rough measure of intensity in the involvement with the scandal to classify politicians into quintiles of parapolitics news coverage. Comparing the two papers suggests that 84% of the politicians are in the same quintile of scandal intensity. Less than 5% of all the candidates that are exposed by *El Tiempo* as parapoliticians are not exposed by *El Espectador*. Similarly, less than 5% of all the candidates that are exposed by *El Espectador* are not exposed by *El Tiempo*. More generally, we conduct a combined Kolmogorov-Smirnov test to investigate whether the distribution of candidates according to their share of parapolitics-related stories is the same across papers. The resulting p-value is near 1, indicating that the null that the two distributions are statistically indistinguishable cannot be rejected.

More formally, in Table 7 we use these classifications of parapoliticians, based on news searches from *El Espectador* (Panel A) and *El Tiempo* (Panel B) –but relying on the overlapping period of available information from the two papers from October 2007 to August 2011– to re-estimate our main regressions, characterizing the electoral success of parapoliticians (as in Table 2). Regardless of the media outlet used to code involvement with the scandal, we corroborate our main result: parapoliticians have a vote share margin advantage relative to non-parapoliticians, which is exacerbated in places where paramilitary presence makes coercion cheaper (Columns 1, 2, 5 and 6)

and attenuated where state capacity makes it more costly (Columns 3, 4, 5 and 6, the latter not statistically significant).

This suggests that it is unlikely that our results which focus on *El Tiempo*'s news archive are biased. Indeed, the size of the estimated coefficients on Panels A and B of Table 7 are extremely similar. In fact, if anything, *El Tiempo underestimates* the consequences of exposure. Table C-6 of Appendix C corroborates these findings.

6 Conclusions

A widespread belief that has been confirmed by recent research is that, by uncovering incompetent or corrupt politicians, free mass media can enhance political accountability and help voters make better decisions. When considering weaker overall institutional environments, most of the literature has focused on situations in which the media is not free, but rather it is captured by political or economic interest groups. This impedes the free flow of unbiased information that Thomas Jefferson had wished for. However, even when weak institutions predominate, the media sometimes still plays an active and unbiased role of providing useful information about political corruption or other wrongdoings by politicians.

In this paper, we argue that in such circumstances the disinfectant role of media can be limited and even have unintended negative consequences. An uneven democracy in which vote casting is not completely free may provide the incentives for certain type of politicians to concentrate their constituencies in places in which coercion or fraud is easier to undertake, even more so if exposed and demonized by mass media. We explored this *dark side* of an active mass media. First, by laying out a simple political economy model that features free media and politicians who compete in elections but might react manipulating the elections. Second, by offering empirical evidence from the Colombian Senate elections from 2002 to 2010.

We derive testable implications from our model, namely that a media scandal that exposes tainted candidates increases coercion to offset the media-driven negative popularity shock. Hence, on top of the direct cost of coercion, an unintended consequence of the media scandal is that exposed candidates may not even see a reduction in their vote share. An additional negative consequence is that the attention from politicians to voters in coerced areas decreases, as less voters respond to policy platforms in these areas.

Our empirical results focus on one salient episode of the recent political history of Colombia, the so called 'parapolitics' scandal, that featured national and local politicians colluding with illegal armed paramilitary groups to obtain votes by exerting violent coercion in areas of little state presence. The scandal was uncovered by the mass media and led to judicial investigations with some involved politicians being sent to jail.

However, our findings suggest it also had the unintended negative consequences that our theory

predicts. Incumbent candidates exposed by the national press to allegedly have ties with paramilitaries, are not only more successful than non-exposed politicians (in terms of their vote share), but also more likely to get their votes in areas controlled by the armed group (but not in areas controlled by other armed groups) and less likely to obtain votes in areas with higher state capacity, as measured by the presence of law enforcement institutions. Perhaps more interesting (and harder to reconcile with potential alternative explanations), we provided evidence on the effects of the media scandal by showing that within the set of candidates exposed by the mass media as allegedly having ties with the paramilitaries, those exposed *before* elections are also more likely to get their votes from paramilitary-controlled regions. We also found some, though weaker, evidence that they are less likely to get them from places with a stronger state. Also, this rearrangement in vote distribution seems to counteract the effects of the media scandal, as these set of candidates fare no worse in elections on average.

To complement Brandeis famous remark, sunlight may well be the best of disinfectants, but not when one can hide in the shade. The institutions for a functioning democracy truly complement each other.

References

- Acemoglu, D., & Robinson, J. A. (2008, March). Persistence of power, elites, and institutions. *American Economic Review*, 98(1), 267-93. Retrieved from <http://ideas.repec.org/a/aea/aecrev/v98y2008i1p267-93.html>
- Acemoglu, D., & Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York, N.Y.: Crown Publishers (Random House).
- Acemoglu, D., Robinson, J. A., & Santos, R. J. (2013). The Monopoly of Violence: Evidence from Colombia. *Journal of the European Economic Association*, 11, 5-44.
- Acevedo, T. (2010). *Monografía político electoral, departamento del cesar: 1997-2007*. Retrieved from http://moe.org.co/home/doc/moe_mre/CD/PDF/cesar.pdf (last accessed 12/5/2012)
- Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., & Zhuravskaya, E. (2015). Radio and the rise of the Nazis in prewar Germany. *The Quarterly Journal of Economics*, 130(4), 1885-1939.
- Andriantsoa, P., Andriasendrarivony, N., Haggblade, S., Minten, B., Rakotojaona, M., Rakotoavy, F., & Razafinimanana, H. (2005). Media proliferation and democratic transition in Africa: The case of Madagascar. *World Development*, 33(11), 1939-1957. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0305750X05001506>
- Asunka, J., Brierley, S., Golden, M., Kramon, E., & Ofosu, G. (2017). Electoral fraud or violence: The effect of observers on party manipulation strategies. *British Journal of Political Science*, 1-23.
- Baron, D. P. (2006). Persistent media bias. *Journal of Public Economics*, 90(1-2), 1-36. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0047272705000216>
- Besley, T., & Burgess, R. (2002). The political economy of government responsiveness: Theory and evidence from India. *The Quarterly Journal of Economics*, 117(4), 1415-1451. Retrieved from <http://ideas.repec.org/a/tpr/qjecon/v117y2002i4p1415-1451.html>
- Besley, T., & Prat, A. (2006). Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review*, 96(3), 720-736. Retrieved from <http://www.atypon-link.com/AEAP/doi/abs/10.1257/aer.96.3.720>
- Brunetti, A., & Weder, B. (2003). A free press is bad news for corruption. *Journal of Public Economics*, 87(7-8), 1801-1824. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0047272701001864>
- Callen, M., & Long, J. D. (2015). Institutional corruption and election fraud: Evidence from a field experiment in afghanistan. *American Economic Review*, 105(1), 354-81.
- Chaves, I., Fergusson, L., & Robinson, J. A. (2015). He who counts elects: Economic elites, political elites, and electoral fraud. *Economics & Politics*, 27(1), 124-159.

- Chong, A., De La O, A. L., Karlan, D., & Wantchekon, L. (2011, December). *Looking beyond the incumbent: The effects of exposing corruption on electoral outcomes* (NBER Working Papers No. 17679). Retrieved from <http://ideas.repec.org/p/nbr/nberwo/17679.html>
- Collier, P., & Vicente, P. C. (2012). Violence, bribery, and fraud: the political economy of elections in sub-saharan africa. *Public choice*, 153(1-2), 117–147.
- Collier, P., & Vicente, P. C. (2014). Votes and violence: Evidence from a field experiment in nigeria. *The Economic Journal*, 124(574).
- Cruz, S. A. G. (2010). *Monografía político electoral, departamento de valle del cauca: 1997-2007*. Retrieved from http://moe.org.co/home/doc/moe_mre/CD/PDF/vdelcauca.pdf (last accessed 12/5/2012)
- Daniele, G., & Dipoppa, G. (2017). Mafia, elections and violence against politicians. *Journal of Public Economics*, 154, 10–33.
- DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., & Zhuravskaya, E. (2011, May). *Unintended media effects in a conflict environment: Serbian radio and croatian nationalism* (NBER Working Papers No. 16989). Retrieved from <http://ideas.repec.org/p/nbr/nberwo/16989.html>
- DellaVigna, S., & Kaplan, E. (2007). The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3), 1187–1234. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/qjec.122.3.1187>
- Djankov, S., & McLiesh, C. (Eds.). (2002). *The right to tell: the role of mass media in economic development*. Washington, D.C.: World Bank Publications. Retrieved from http://books.google.com/books?hl=en&lr=&id=56punueI7G0C&oi=fnd&pg=PP7&dq=The+Right+to+Tell:+the+Role+of+Mass+Media+in+Economic+Development&ots=y0jVsk3nAV&sig=PNWiDshgFtvT_2vNR6ysefz9qI
- Dube, O., & Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from colombia. *The Review of Economic Studies*, 80(4), 1384–1421.
- Durante, R., & Knight, B. (2012). Partisan control, media bias, and viewer responses: Evidence from berlusconis italy. *Journal of the European Economic Association*, 10(3), 451–481.
- Eisensee, T., & Strömberg, D. (2007). News droughts, news floods, and US disaster relief. *The Quarterly Journal of Economics*, 122(2), 693–728. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/qjec.122.2.693>
- Fergusson, L., Querubin, P., Ruiz-Guarin, N. A., & Vargas, J. F. (2017). The real winner’s curse. *Documento CEDE 2017-05*.
- Ferraz, C., & Finan, F. (2008). Exposing corrupt politicians: The effects of Brazil’s publicly released audits on electoral outcomes. *Quarterly Journal of Economics*, 123(2), 747–793.
- Gentzkow, M., & Shapiro, J. (2004). Media, education and anti-americanism in the Muslim world. *Journal of Economic Perspectives*, 18(3), 117–133.

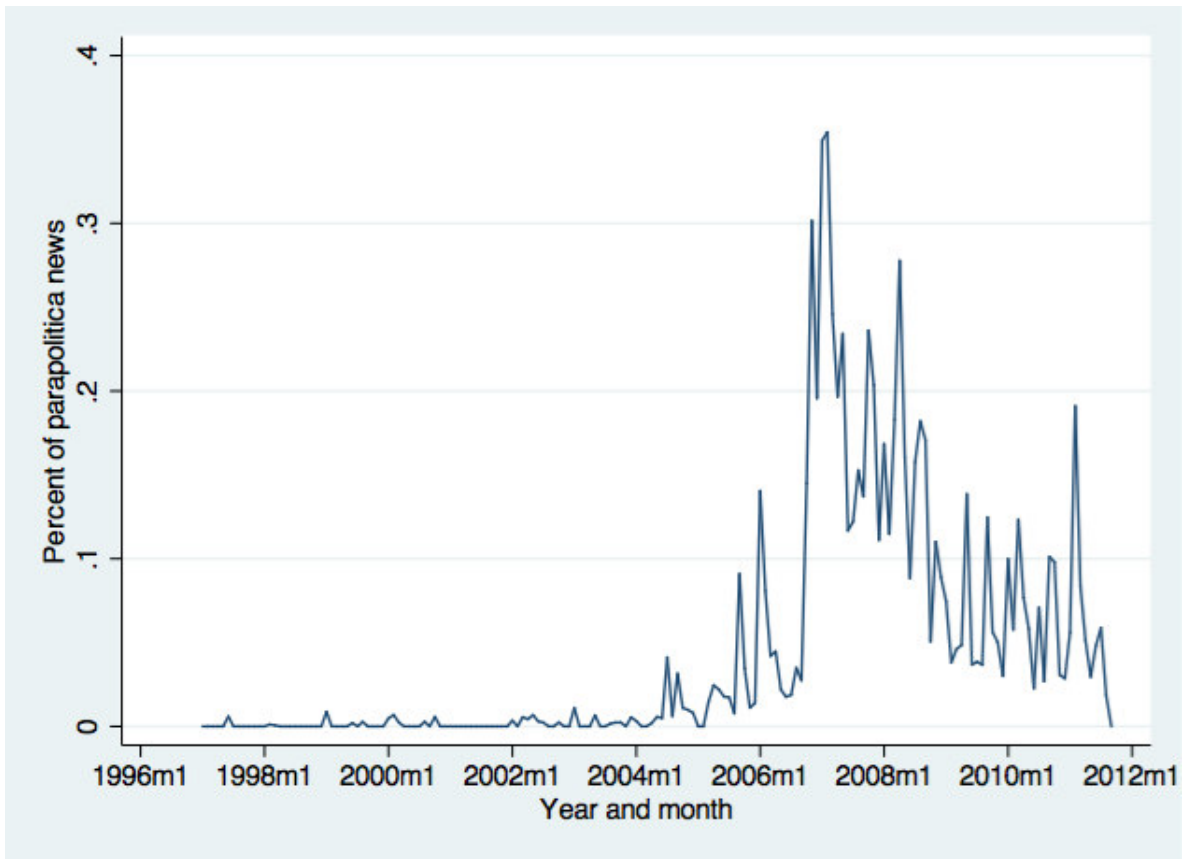
- Gentzkow, M., & Shapiro, J. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280–316. Retrieved from <http://www.journals.uchicago.edu/doi/abs/10.1086/499414>
- Gentzkow, M., & Shapiro, J. (2008). Competition and truth in the market for news. *Journal of Economic Perspectives*, 22(2), 133–154.
- Gentzkow, M., & Shapiro, J. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35–71. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=947640 doi: 10.3982/ECTA7195
- Gentzkow, M., Shapiro, J. M., & Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, 101(7), 2980–3018.
- Glaeser, E. L., Gentzkow, M. A., & Goldin, C. (2004). *The rise of the Fourth Estate: How newspapers became informative and why it mattered*. Cambridge. (NBER Working Paper Series, Working Paper 10791)
- Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/003355505775097542>
- Harish, S., & Little, A. T. (2017). The political violence cycle. *American Political Science Review*, 111(2), 237–255.
- Jetter, M. (2017). The effect of media attention on terrorism. *Journal of Public Economics*, 153, 32–48.
- Laverde, J. D. (2012). “juan carlos martínez: un intocable en desgracia”. *El Espectador*, July. Retrieved from <http://www.elespectador.com/impreso/judicial/articulo-359827-juan-carlos-martinez-un-intocable-desgracia>
- Lindbeck, A., & Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, 52(3), 273–297. Retrieved from <http://search.proquest.com/docview/194651204?accountid=34489>
- López, C. (2007). La Ruta de la Expansión Paramilitar y la Transformación Política de Antioquia. In M. Romero (Ed.), *Para Política: La Ruta de la Expansión Paramilitar y los Acuerdos Políticos*. Bogotá: Corporación Nuevo Arco Iris: Intermedio.
- López, C. (2010). *Y refundaron la patria... de cómo mafiosos y políticos reconfiguraron el estado colombiano*. Bogotá: Debate.
- McMillan, J., & Zoido, P. (2002). How to subvert democracy: Montesinos in Peru. *Journal of Economic Perspectives*, 18(4), 69–92.
- Naidu, S. (2012). *Suffrage, schooling, and sorting in the post-bellum U.S. south* (Working Paper No. 18129). National Bureau of Economic Research. (<http://www.nber.org/papers/w18129.pdf>)
- Petrova, M. (2008). Inequality and media capture. *Journal of Public Economics*, 92(1–2), 183–212.

- Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0047272707000606>
- Prat, A., & Strömberg, D. (2005). *Commercial television and voter information*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=772002 (CEPR Discussion Paper 4989)
- Prat, A., & Strömberg, D. (2013). The political economy of mass media. In D. Acemoglu, M. Arellano, & E. Dekel (Eds.), *Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress*. Cambridge University Press.
- Puglisi, R. (2011). Being the new york times: the political behaviour of a newspaper. *The BE Journal of Economic Analysis & Policy*, 11(1).
- Restrepo, J. (2008). “la parapolítica y la prensa”. *Revista Cambio*, May. Retrieved from http://www.cambio.com.co/portadacambio/778/ARTICULO-PRINTER_FRIENDLY-PRINTER_FRIENDLY_CAMBIO-4214207.html
- Restrepo, J., Spagat, M., & Vargas, J. (2004). The dynamics of the columbian civil conflict: A new dataset. *Homo Oeconomicus*, 21, 396-429. Retrieved from <http://ideas.repec.org/a/hom/homoec/v21y2004p396-429.html>
- Robinson, J. A., & Torvik, R. (2009). The real swing voter’s curse. *American Economic Review*, 99(2), 310–15.
- Roland, G., & Zapata, J. G. (2004). Colombia’s electoral and party system: Proposals for reform. In A. Alesina (Ed.), *Institutional reforms in Colombia*. Cambridge: MIT Press.
- Semana. (2006a). “cómo se hizo el fraude”. *Revista Semana*, April. Retrieved from <http://www.semana.com/nacion/como-hizo-fraude/93809-3.aspx>
- Semana. (2006b). “el senador álvaro aráujo confiesa que estuvo en una fiesta con ‘jorge 40’ antes de su desmovilización”. *Revista Semana*, November. Retrieved from <http://www.semana.com/on-line/senador-alvaro-araujo-confiesa-estuvo-fiesta-jorge-40-antes-su-desmovilizacion/98224-3.aspx>
- Semana. (2007). “¿Quién es el Fidel Cano que hoy dirige ?El Espectador? que otro Fidel Cano fundó hace 120 años?”. *Revista Semana*, March. Retrieved from <http://www.semana.com/nacion/quien-fidel-cano-hoy-dirige-espectador-otro-fidel-cano-fundo-hace-120-anos/101595-3.aspx>
- Snyder, J. M., & Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, 118(2), 355-408. Retrieved from <http://ideas.repec.org/a/ucp/jpolec/v118y2010i2p355-408.html>
- Strömberg, D. (2004a). Mass media competition, political competition, and public policy. *Review of Economic Studies*, 71(1), 265–284. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=515533
- Strömberg, D. (2004b). Radio’s impact on public spending. *Quarterly Journal of Economics*, 119(1), 189–221. Retrieved from <http://www.mitpressjournals.org/doi/abs/10.1162/>

003355304772839560

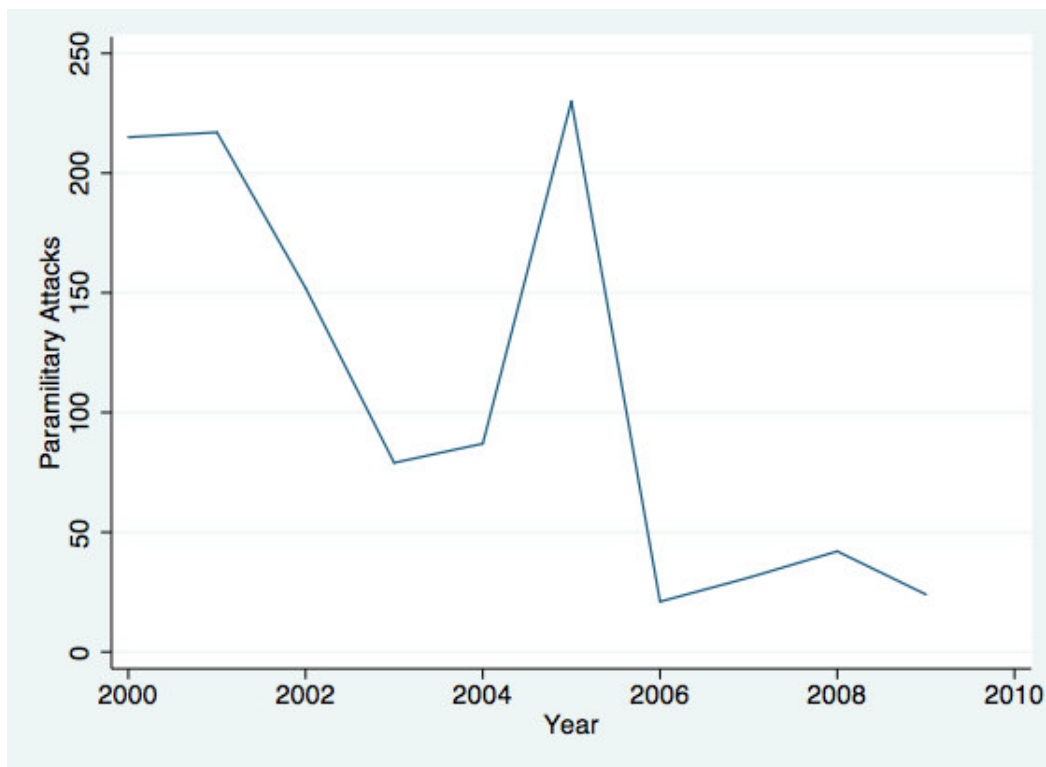
Valencia, L. (2007). Los Caminos de la Alianza entre Los Paramilitares y los Políticos. In M. Romero (Ed.), *Para Política: La Ruta de la Expansión Paramilitar y los Acuerdos Políticos*. Bogotá: Corporación Nuevo Arco Iris: Intermedio.

Figure 1: Percent of parapolitics news stories,
out of all recorded stories about Congress candidates



Authors' coding from *El Tiempo* online archive, see Appendix section B for details. The sample of politicians included are all candidates to the House or Senate elections with previous experience either in the House or Senate.

Figure 2: Evolution of paramilitary attacks: 2000-2009



Source: Authors' update of Restrepo, Spagat, and Vargas (2004).

Table 1A: Descriptive statistics – Municipal level variables

	mean	std. dev.	min	max
Candidate's municipality vote share	0.47	2.68	0	95.66
Paramilitary presence	0.007	1.16	-0.65	11.51
Dummy for high paramilitary presence	0.403	0.49	0	1
State capacity index	0.0002	1.7	-0.24	7.47
Dummy for high state capacity	0.26	0.43	0	1
Dummy for municipality captured by guerrilla	0.03	0.17	0	1
Dummy for high guerrilla attacks per capita	0.40	0.49	0	1
Dummy for unusually high turnout	0.40	0.49	0	1

Notes: For sources and variable definitions see Section 3.

Table 1B: Descriptive statistics – Candidate level variables

	All politicians				Parapoliticians		Non-parapoliticians	
	mean (1)	std. dev. (2)	min (3)	max (4)	mean (5)	std. dev. (6)	mean (7)	std. dev. (8)
<i>Panel A: Involvement with parapolitics scandal</i>								
Parapolitician Dummy	0.33	0.47	0	1	1	0	0	0
Perc. parapolitics news	0.06	.14	0	0.86	0.17	0.2	0	0
Dummy news before elections	0.17	0.38	0	1	0.55	0.5	0	0
<i>Panel B: Candidate level outcomes and controls</i>								
Average News in <i>El Tiempo</i>	361.9	624.88	1	4769	355.13	336.78	365.29	728.44
Male	0.88	0.32	0	1	0.87	0.34	0.89	0.31
College Degree	0.9	0.3	0	1	0.89	0.31	0.91	0.29
Replaced Candidate	0.04	0.2	0	1	0.04	0.2	0.04	0.2
Average Years in Politics	13.18	8.71	2	46	12.29	8.18	13.63	8.94
Average Periods in Congress	2.1	1.37	1	11	2.08	1.52	2.12	1.29

Notes: Source for Panel A is authors' coding from *El Tiempo* online archive, see Appendix section B for details. The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel B, except for candidate vote share which is from Official Sources (*Registraduría Nacional del Estado Civil*), candidate controls are from the NGO *Congreso Visible*.

Table 2: Parapoliticians' relative success in shade vs sunlight areas

Dependent variable: Candidate's municipal vote share						
	Shade interaction		Sunlight interaction		Both interactions	
	Paramilitary presence (1)	Dummy high paramilitary presence (2)	State capacity (3)	Dummy high state capacity (4)	Continuous measures (5)	Dummies (6)
<i>Panel A: Extensive margin exposure</i>						
Parapolitician Dummy	0.119 (0.015)*** [0.046]**	0.058 (0.020)*** [0.059]	0.119 (0.015)*** [0.046]**	0.123 (0.015)*** [0.048]**	0.119 (0.015)*** [0.046]**	0.061 (0.021)*** [0.062]
Parapolitician Dummy x Municipal shade	0.036 (0.011)*** [0.023]	0.152 (0.029)*** [0.072]**			0.035 (0.011)*** [0.023]	0.150 (0.030)*** [0.072]**
Parapolitician Dummy x Municipal sunlight			-0.008 (0.002)*** [0.004]**	-0.075 (0.043)* [0.052]	-0.007 (0.002)*** [0.004]**	-0.043 (0.043) [0.058]
<i>Panel B: Intensive margin exposure</i>						
Perc. parapolitics news	0.438 (0.067)*** [0.126]***	0.151 (0.087)* [0.156]	0.437 (0.068)*** [0.127]***	0.456 (0.071)*** [0.131]**	0.439 (0.067)*** [0.126]***	0.166 (0.091)* [0.165]
Perc. parapolitics news x Municipal shade	0.254 (0.064)*** [0.104]**	0.784 (0.134)*** [0.316]**			0.252 (0.064)*** [0.104]**	0.717 (0.135)*** [0.323]**
Perc. parapolitics news x Municipal sunlight			-0.034 (0.011)*** [0.012]***	-0.347 (0.204)* [0.198]*	-0.027 (0.010)*** [0.011]**	-0.203 (0.204) [0.213]
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	337,841	337,841	337,841	337,841	337,841	337,841
Number of candidates	254	254	254	254	254	254

Notes: Robust standard errors, clustered at the municipality level (in parentheses), and two-level clustered at the municipality and candidate level [in brackets]. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Dependent variable is the vote share of each candidate in each municipality/electoral year. The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel A, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In Panel B, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. All regressions control for (the log of) municipal population. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

Table 3: Comparability of parapoliticians exposed before and after the elections

Dependent variable: Dummy for exposure before elections		
	(1)	(2)
Replaced Candidate	0.0252 (0.129)	0.0544 (0.114)
College Degree	-0.1000 (0.0928)	-0.108 (0.0931)
Male	0.146 (0.0988)	0.155 (0.100)
Change of House	0.0431 (0.0710)	0.0264 (0.0726)
Average Years in Politics	-0.0007 (0.0060)	0.0006 (0.0063)
Average Periods in Congress	0.0381* (0.0211)	0.0366* (0.0217)
Total news stories in <i>El Tiempo</i>	0.0002** (0.0001)	
Non-paramilitary news stories in <i>El Tiempo</i>		0.0001 (0.0001)
R-squared	0.803	0.797
Time fixed effects	Yes	Yes
Observations	119	119
Number of candidates	77	77

Notes: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Dependent variable is a dummy that equals 1 if the politician is exposed *before* the corresponding election. Sample includes only the set of parapoliticians (politicians exposed as having links with paramilitaries at any point in time).

Table 4: Relative success in shade vs sunlight areas of parapoliticians exposed before elections relative to parapoliticians exposed after

Dependent variable: Candidate's municipal vote share						
	Shade interaction		Sunlight interaction		Both Interactions	
	Paramilitary presence	Dummy high paramilitary presence	State capacity	Dummy high state capacity	Continuous measures	Dummies
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy news before elections	-0.020 (0.029) [0.082]	-0.021 (0.032) [0.085]	-0.020 (0.030) [0.082]	-0.024 (0.030) [0.083]	-0.021 (0.029) [0.082]	-0.025 (0.033) [0.088]
Dummy news before elections x Municipal shade	0.040 (0.012)*** [0.026]	0.002 (0.038) [0.081]			0.040 (0.012)*** [0.026]	0.005 (0.038) [0.082]
Dummy news before elections x Municipal sunlight			0.004 (0.005) [0.006]	0.056 (0.043) [0.047]	0.005 (0.005) [0.007]	0.046 (0.040) [0.047]
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112,269	112,269	112,269	112,269	112,269	112,269
Number of candidates	77	77	77	77	77	77

Notes: Robust standard errors, clustered at the municipality level (in parentheses), and two-level clustered at the municipality and candidate level [in brackets]. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Dependent variable is the vote share of each candidate in each municipality/electoral year. The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. 'News before elections' is a dummy variable that equals 1 if the candidate was linked to paramilitaries before the corresponding election. All regressions control for (the log of) municipal population. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

Table 5: Parapoliticians relative success in high turnout areas

Dependent variable: Candidate's municipal vote share			
Exposure measure:	Parapolitician dummy (1)	% Parapolitics news (2)	Dummy news before elections (3)
Exposure	0.067 (0.015)*** [0.050]	0.261 (0.067)*** [0.169]	-0.035 (0.032) [0.084]
Exposure \times Dummy unusually high turnout	0.132 (0.029)*** [0.068]*	0.438 (0.133)*** [0.308]	0.037 (0.034) [0.085]
Municipality fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes
Observations	337,643	337,643	112,211
Number of candidates	254	254	77

Notes: Robust standard errors, clustered at the municipality level (in parentheses), and two-level clustered at the municipality and candidate level [in brackets]. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Dependent variable is the vote share of each candidate in each municipality/electoral year. The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In column 1, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In column 2, '% parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. In column 3, 'News before elections' is a dummy variable that equals 1 if the candidate was linked to paramilitaries before the corresponding election. All regressions control for (the log of) municipal population. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

Table 6: Placebo: Parapoliticians' relative success in guerrilla areas

Dependent variable:	Captured by guerrilla (1)	High guerr. attacks (2)
<i>Panel A. Extensive margin exposure</i>		
Parapolitician Dummy	0.065 (0.021)*** [0.060]	0.077 (0.023)*** [0.061]
Parapolitician Dummy \times Guerr. interaction	-0.085 (0.039)** [0.172]	-0.066 (0.031)** [0.061]
Parapolitician Dummy \times Dummy high param. presence	0.147 (0.030)*** [0.072]**	0.178 (0.031)*** [0.074]**
<i>Panel B. Intensive margin exposure</i>		
Perc. parapolitics news	0.187 (0.091)** [0.163]	0.238 (0.101)** [0.174]
Perc. parapolitics news \times Guerr. interaction	-0.956 (0.121)*** [0.256]***	-0.392 (0.129)*** [0.190]**
Perc. parapolitics news \times Dummy high param. presence	0.715 (0.136)*** [0.324]**	0.902 (0.137)*** [0.316]***
<i>Panel C. Exposure before election</i>		
Dummy news before elections	-0.013 (0.032) [0.086]	-0.017 (0.033) [0.086]
Dummy news before elections \times Guerr. interaction	-0.243 (0.057)*** [0.222]	-0.021 (0.043) [0.073]
Dummy news before elections \times Dummy high param. presence	0.003 (0.038) [0.082]	0.014 (0.044) [0.085]
Time fixed effects	Yes	Yes
Candidate controls	Yes	Yes
Observations (Panels A & B)	331,701	331,701
Observations (Panel C)	110,289	110,289
Number of candidates (Panels A & B)	254	254
Number of candidates (PanelC)	77	77

Notes: Robust standard errors, clustered at the municipality level (in parentheses), and two-level clustered at the municipality and candidate level [in brackets]. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Dependent variable is the vote share of each candidate in each municipality/electoral year. The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel A, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In Panel B, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. In Panel C, 'News before elections' is a dummy variable that equals 1 if the candidate was linked to paramilitaries before the corresponding election. All regressions control for (the log of) municipal population. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

Table 7: Verifying media bias: Parapoliticians' relative performance using *El Espectador*

Dependent variable:	Shade interaction		Sunlight interaction		Both interactions	
	Paramilitary presence (1)	Dummy high paramilitary attacks (2)	State capacity (3)	Dummy high state capacity (4)	Continuous measures (5)	Dummies (6)
<i>Panel A. Exposure based on El Espectador</i>						
Parapolitician Dummy	0.139 (0.017)*** [0.047]***	0.055 (0.022)** [0.062]	0.139 (0.017)*** [0.047]***	0.144 (0.018)*** [0.049]***	0.139 (0.017)*** [0.047]***	0.059 (0.024)** [0.065]
Parapolitician Dummy × Municipal shade	0.042 (0.015)*** [0.026]	0.208 (0.034)*** [0.080]**			0.042 (0.015)*** [0.026]	0.206 (0.034)*** [0.082]**
Parapolitician Dummy × Municipal sunlight			-0.010 (0.003)*** [0.004]***	-0.098 (0.051)* [0.059]*	-0.008 (0.003)*** [0.004]**	-0.055 (0.051) [0.065]
<i>Panel B. Exposure based on El Tiempo</i>						
Parapolitician Dummy	0.136 (0.016)*** [0.047]***	0.065 (0.022)*** [0.060]	0.136 (0.016)*** [0.047]***	0.141 (0.017)*** [0.048]***	0.136 (0.016)*** [0.047]***	0.068 (0.023)*** [0.063]
Parapolitician Dummy × Municipal shade	0.043 (0.012)*** [0.025]*	0.177 (0.032)*** [0.077]**			0.043 (0.012)*** [0.025]*	0.174 (0.032)*** [0.078]**
Parapolitician Dummy × Municipal sunlight			-0.009 (0.003)*** [0.004]**	-0.090 (0.045)** [0.054]*	-0.008 (0.002)*** [0.004]**	-0.053 (0.045) [0.060]
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	321,557	321,557	321,557	321,557	321,557	321,557
Number of candidates	244	244	244	244	244	244

Notes: Robust standard errors, clustered at the municipality level (in parentheses), and two-level clustered at the municipality and candidate level [in brackets]. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Dependent variable is the vote share of each candidate in each municipality/electoral year. The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel A, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In Panel B, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. All regressions control for (the log of) municipal population. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

A Appendix. Model Extension: Media Scandal and Coercion Costs

The endogenous increase in coercion after a media scandal is the main driver of our key theoretical results. However, in our baseline model we ignored that more media exposure can also make coercion more costly, discouraging it and potentially eliminating the main mechanisms we highlighted in the main text. In this section, we extend our baseline model in such direction and show that, while this indeed attenuates the increase in coercion, our main results are still present under admissible parameter values.

Let the cost of coercion be given by,

$$C(\{\lambda_m\}_{m=1}^N) = \sum_{m=1}^N \psi_m \mu_m^\alpha c(\lambda_m), \quad (\text{A-1})$$

where $\alpha \geq 0$. The baseline model corresponds to the special case with $\alpha = 0$.

Following the same steps as before, the first order conditions of politician A indicate that equilibrium coercion λ_m^* is now given by:

$$\lambda_m^* = \max \left\{ 0, \frac{1}{\psi_m \mu_m^\alpha} \left(\frac{1}{2} + \phi_m (\mu_m - \sigma_m) \right) \right\},$$

Hence, we can write

$$\frac{\partial \lambda_m^*}{\partial \mu_m} = \frac{\phi_m}{\psi_m \mu_m^\alpha} - \alpha \frac{\lambda_m^*}{\mu_m}.$$

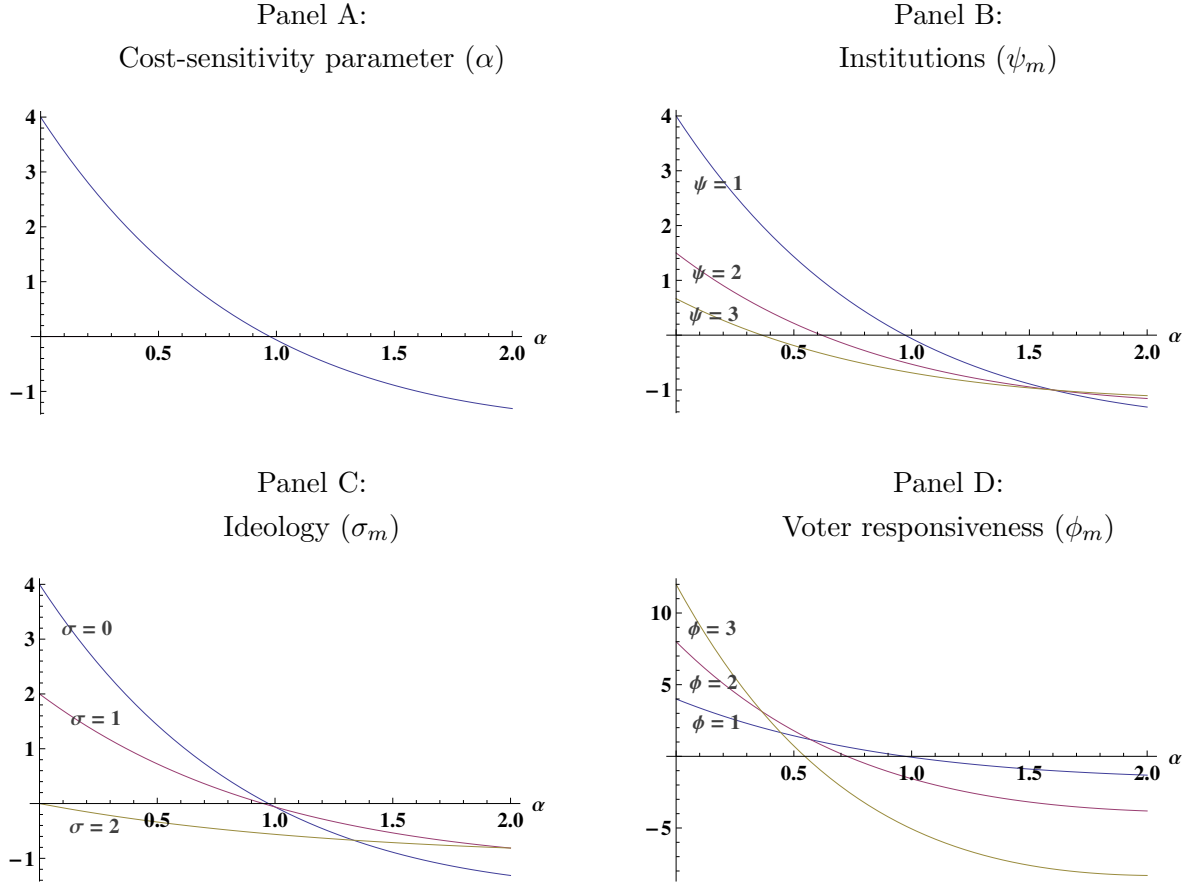
After some algebra, this implies that our main object of interest, the change in parapolitician support in each municipality after the media scandal, is given by:

$$\frac{d\pi_m^A}{d\mu_m} = -(1 - 2\lambda_m^*)\phi_m - \alpha(\lambda_m^*)^2 \psi_m \mu_m^{\alpha-1} \quad (\text{A-2})$$

While in this extension there is no simple closed-form parameter restriction that establishes the conditions under which $\frac{d\pi_m^A}{d\mu_m} \geq 0$, the graphical analysis that follows shows that the endogenous response of coercion may still imply that electoral support for politician A does not fall after the media scandal. Also, we show that the main comparative static predictions of the baseline model still hold.

We set the following baseline values for the model parameters: $\phi_m = \psi_m = 1$, $\mu_m = 2$ and $\sigma_m = 0$. Panel A in Figure A-1 shows that $\frac{d\pi_m^A}{d\mu_m} > 0$ for sufficiently low α . That is, as long as coercion costs do not increase too sharply with the media scandal, the main result in the baseline model still holds under admissible parameter values: given the endogenous response of coercion, the media scandal may not hurt the parapolitician's total electoral support.

Figure A-1: Change in Politician Electoral Support with Media Scandal
 $(d\pi_m^A/d\mu_m)$



Notes: Each panel graphs $d\pi_m^A/d\mu_m$, the change in total electoral support to candidate A (the parapolitician) as a result of a media scandal in municipality m . The algebraic expression is given in (A-2). This derivative is plotted as a function of the cost-sensitivity parameter, α , capturing the extent to which a media scandal increases coercion costs in a given municipality. From (A-1), the cost of coercion λ_m in each municipality is given by $\psi_m \mu_m^\alpha \lambda^2 / 2$. Baseline values for the model parameters are $\phi_m = \psi_m = 1$, $\mu_m = 2$ and $\sigma_m = 0$. Panel B checks sensitivity of $d\pi_m^A/d\mu_m$ to different values of the quality of institutions (ψ_m), Panel C to ideology (σ_m), and Panel D to voter responsiveness (ϕ_m).

Panels B-D show the main comparative static implications of the model. In Panel B, we plot $d\pi_m^A/d\mu_m > 0$ for the baseline value $\psi_m = 1$ for the quality of local institutions, together with $\psi_m = 2$ and $\psi_m = 3$. Clearly, an increase in the quality of institutions makes it less likely that the parapolitician may compensate for his loss in popularity through an increase in coercion. In Panel C, we allow voter ideology σ_m to increase from its baseline value of 0 to 2. Recall that for the purpose of our

example we set $\mu_m = 2$, meaning that the key value of σ_m is 2. Indeed, for $\sigma_m > 2$ the net average partisanship of voters ($\sigma_m - \mu_m$) favors politician A . Thus, free voters are on average supporting politician A , which reduces the returns to increasing coercion. In line with this intuition, Panel C shows that starting at $\sigma_m = 2$, the electoral support for the parapolitician falls with the media scandal even if exposure does not make coercion more costly²⁹. Finally, Panel D explores changes in voter responsiveness, from the baseline value of $\sigma = 1$ to $\sigma = 3$. We observe that as voters become more responsive, parapolitician support increases after the scandal for lower values of α . Intuitively, when voters are more responsive many votes are lost as a result of the media scandal, and it is less likely that the parapolitician may compensate for the fall in popularity with coercion. Notice, however, that when they do compensate for the fall in popularity and $d\pi_m^A/d\mu_m > 0$, the increase is larger for large ϕ_m , as more votes are prevented from going to the opponent through coercion in this case.

²⁹Indeed, recall from the baseline model that $\frac{d\pi_m^A}{d\mu_m} \gtrless 0$ if $\frac{1}{2} + \phi_m (\mu_m - \sigma_m) \gtrless \frac{1}{2} \psi_m$. Hence, for every $\mu_m - \sigma_m < 0$ we have $\frac{d\pi_m^A}{d\mu_m} < 0$ even if coercion costs are independent of the media scandal.

B Appendix. Coding of Politician News Stories

Our data set is event-based and relies on information on every news story (and op-ed) published from January 1st 1997 to August 1st 2011 in *El Tiempo*, Colombia’s largest newspaper with nationwide circulation and coverage. We search for stories in every section of the newspaper about each of the incumbent congressmen in this 15-year period. For every story we record the date, the newspaper section in which it appears (and if it appears in more than one section), and whether it links the politician with parapolitics activity, other corruption practices or an electoral crime (like fraud or vote-buying).

The first phase of the data gathering took place from August to October 2010. A team of over 30 undergraduate students at Universidad del Rosario were trained to use the *El Tiempo* electronic archive and code stories about each candidate. We searched for all possible variations of the (frequently two) names and two last names of each candidate. The heading of each story (or the whole story if necessary) was checked to discard homonyms and corroborate it referred to the politician.

We double-coded the set of news for a fourth of all the candidates in the sample (the rest of the candidate sample were coded again in the second phase of the data gathering). This was done by different coders in different coding sessions taking place in different days, to ensure the two codings of the same candidate were independent. Where we found large and systematic differences between the two codings, the candidate was again coded twice in the second phase³⁰. In this first phase, coders were paid according to their productivity, as measured both by the amount of candidate-stories coded per session, as well as the average coding quality.

The second phase of the data gathering was part of an economic experiment conducted by Mariana Blanco and Patricio Dalton between August and September of 2011. During this phase, we completed the double checking for the candidates of the first phase, coded the *El Tiempo* news stories between July 2010 and August 2011 for of all candidates, and coded stories from *El Espectador*³¹, coding these twice for each candidate. In the experiment, 200 undergraduate students from Universidad del Rosario and Universidad de los Andes were hired for the codification and divided in 12 groups (half doing news from *El Tiempo* and the other half doing news from *El Espectador*) and the same instructions from the first phase were given to the students³².

In each session the coders received a number of random candidates with the information of the start date to code. In the experiment, students had the incentive to do a good job because they were likely to be “fired” and do instead other activity with lower salary or were paid a smaller amount for just attending the session without working. From the first day of work and in each

³⁰ All coders were double-coded by a random peer in at least one of the politicians assigned.

³¹ *El Espectador* only became a daily newspaper since October of 2007 and only since then are its news stories available in the online archive

³² The instructions for *El Espectador* codification differed in some minor aspects because of differences in its website.

session, the coders within each group were classified according to their productivity and quality and in the next session only the more productive ones codified news stories. The number of students working varied in each session and fired coders had the chance to be hired again. Thus, checking quality was of extreme importance for the experiment and after each session at least 15% of the news stories coded by each student were checked. Also, codifications by any student in one session with more than 40% of errors were excluded altogether. With all candidates coded twice in both *El Tiempo* and *El Espectador* newspaper, this allowed us to cross-check the quality of the coding.

C Appendix. Interpreting ‘Exposure’ in candidate-level specifications

Refer to coefficient $\hat{\beta}_1$ in regression (11), and recall our theoretical model. Notice that a regression of $\overline{\text{Shade}}_c$ (and similarly for $\overline{\text{Sunlight}}_c$) on a parapolitician dummy simply uncovers the difference in $\overline{\text{Shade}}_c$ for a parapolitician and a non-parapolitician. That is, in terms of our model, $\beta_1 = \overline{\text{Shade}}_A - \overline{\text{Shade}}_B$. Thus, from the definition of $\overline{\text{Shade}}_c$ in (10) and writing the vote shares obtained by each candidate in each municipality using the model notation, we have:

$$\hat{\beta}_1 = \overline{\text{Shade}}_A - \overline{\text{Shade}}_B = \sum_m \left(\frac{\pi_m^A}{\Pi^A} \right) \text{Shade}_m - \sum_m \left(\frac{1 - \pi_m^A}{N - \Pi^A} \right) \text{Shade}_m$$

Let $\overline{\text{Shade}}_m = \sum_m \text{Shade}_m / N$ be the average of a specific ‘shade’ or ‘sunlight’ characteristic across municipalities, and $\bar{\pi}_m^A = \sum_m \pi_m^A / N = \Pi^A / N$ the vote share for A across municipalities. Then, using $\text{Cov} \{ \text{Shade}_m, \pi_m^A \} = (1/N) \sum_m \text{Shade}_m \pi_m^A - \overline{\text{Shade}}_m \bar{\pi}_m^A$, we can rewrite the above the expression after some algebra as:

$$\hat{\beta}_1 = \frac{\text{Cov} \{ \text{Shade}_m, \pi_m^A \}}{\bar{\pi}^A (1 - \bar{\pi}^A)}. \quad (\text{A-3})$$

The key implication of (A-3) is that the expected sign of $\hat{\beta}_1$ is governed by the expected correlation between each municipal ‘shade’ (or ‘sunlight’) characteristic and parapolitician support π_m^A . Thus, for instance, since we know that all else equal parapoliticians are able to exert more coercion and thus get more votes in places with paramilitaries, $\text{Cov} \{ \text{Shade}_m, \pi_m^A \} > 0$. This logic also underlines our expectations regarding the effects of the media scandal. For example, since following exposure of the parapolitician by the media we expect him to react by increasing coercion especially in places with paramilitaries, we also have $\text{Cov} \{ \text{Shade}_m, \pi_m^A \} > 0$ in regressions that include only the set of parapoliticians and compare those exposed before elections with those exposed after.

C.1 Results

Table C-1: Descriptive statistics – Municipal characteristics weighted by municipal candidate vote share

	All politicians				Parapoliticians		Non-parapoliticians	
	mean (1)	std. dev. (2)	min (3)	max (4)	mean (5)	std. dev. (6)	mean (7)	std. dev. (8)
Candidate's national vote share	0.41	0.28	0.01	2.09	0.466	0.225	0.378	0.29
Paramilitary presence	2.79	2.05	0.46	22.21	3.45	2.53	2.46	1.66
Dummy high paramilitary presence	22.2	14.4	2.1	81.430	27.6	16.92	19.47	12.09
State capacity index	19.03	15.97	1.26	79.07	12.83	11.89	22.18	16.85
Dummy high state capacity	51.49	17.66	9.81	90.72	44.13	15.76	55.18	17.44
Dummy municipality captured by guerrilla	2.28	6.23	0.44	54.65	1.99	6.1400	2.43	6.28
Dummy high guerrilla attacks pc	4.23	4.61	0.028	27.57	3.72	4.07	4.48	4.85
Dummy unusually high turnout	24.52	19.03	1.1	96.72	32.59	22.77	20.47	15.37

Notes: The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. For each candidate c and time period t we compute $\bar{y}_{ct} = \sum_m \pi_{mt}^c y_m$, for municipality characteristic y_m and with π_{mt}^c , the vote share of candidate c (out of his own total votes) in municipality m at time t . We rescale y_m to vary between 0 and 100. In particular, we compute y_m as $y_m = 100 \frac{\tilde{y}_m - \min_m \tilde{y}_m}{\max_m \tilde{y}_m - \min_m \tilde{y}_m}$, where \tilde{y}_m is the unscaled municipality characteristic.

Table C-2: Parapoliticians' vote share and vote concentration in shade vs sunlight areas – Candidate level regressions

Dependent variable:	Performance	Shade		Sunlight	
	Vote share (1)	Paramilitary presence (2)	Dummy high paramilitary presence (3)	State capacity (4)	Dummy high state capacity (5)
<i>Panel A. Extensive margin exposure</i>					
Parapolitician Dummy	0.10*** (0.02)	1.056*** (0.266)	8.606*** (1.739)	-9.028*** (1.501)	-10.94*** (1.794)
Constant	0.30*** (0.06)	2.016*** (0.513)	17.15*** (4.004)	25.51*** (3.762)	65.53*** (4.534)
R-squared	0.331	0.091	0.100	0.185	0.196
<i>Panel B. Intensive margin exposure</i>					
Perc. parapolitics news	0.39*** (0.09)	4.625*** (0.896)	33.37*** (7.228)	-29.83*** (3.232)	-38.86*** (5.392)
Constant	0.31*** (0.06)	2.044*** (0.459)	17.75*** (3.870)	24.43*** (3.714)	64.04*** (4.436)
R-squared	0.339	0.129	0.123	0.181	0.204
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes	Yes	Yes
Observations	356	356	356	356	356
Number of candidates	254	254	254	254	254

Notes: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Each dependent variable in columns 2-5 is the municipal (shade or sunlight) characteristic in the column header weighted by the candidate's municipal vote share each electoral year (see equation 10). The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel A, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In Panel B, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

Table C-3: Vote share and vote concentration in shade vs sunlight areas of parapoliticians exposed before elections relative to parapoliticians exposed after – Candidate level regressions

Dependent variable:	Performance	Shade		Sunlight	
	Vote share (1)	Paramilitary presence (2)	Dummy high paramilitary presence (3)	State capacity (4)	Dummy high state capacity (5)
Dummy news before elections	0.06 (0.05)	1.065** (0.513)	6.748 (4.160)	-4.239* (2.402)	-3.341 (3.620)
Constant	0.60*** (0.09)	5.039*** (1.011)	40.66*** (8.374)	12.23*** (4.236)	42.29*** (7.176)
R-squared	0.282	0.238	0.236	0.293	0.244
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes	Yes	Yes
Observations	119	119	119	119	119
Number of candidates	77	77	77	77	77

Notes: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Each dependent variable in columns 2-5 is the municipal (shade or sunlight) characteristic in the column header weighted by the candidate's municipal vote share each electoral year (see equation 10). The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. 'Dummy news before elections' is a dummy variable that equals 1 if the candidate was linked to paramilitaries before the corresponding election. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

**Table C-4: Parapoliticians vote concentration in high turnout areas –
Candidate-level regressions**

Dependent variable: Support in high turnout areas			
Exposure measure:	Parapolitician dummy (1)	% Parapolitics news (2)	Dummy news before elections (3)
Exposure	12.37*** (2.259)	39.44*** (8.757)	11.04* (6.228)
Constant	19.75*** (4.497)	21.36*** (4.467)	51.26*** (11.15)
R-squared	0.154	0.143	0.159
Time fixed effects	Yes	Yes	Yes
Candidate controls	Yes	Yes	Yes
Observations	356	356	119
Number of candidates	254	254	77

Notes: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. The dependent variable aggregates a municipal high turnout dummy weighting by the candidate's municipal vote share each electoral year (see equation 10). The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In column 1, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In column 2, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. In column 3, 'Dummy news before elections' is a dummy variable that equals 1 if the candidate was linked to paramilitaries before the corresponding election. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

**Table C-5: Pabebo: Parapoliticians' vote concentration in guerrilla areas –
Candidate-level regressions**

Dependent variable:	Captured by guerrilla (1)	High guerr. attacks (2)
<i>Panel A. Extensive margin exposure</i>		
Parapolitician Dummy	-0.255 (0.668)	-0.728 (0.492)
Constant	-0.779 (0.766)	0.848 (0.911)
R-squared	0.029	0.048
<i>Panel B. Intensive margin exposure</i>		
Perc. parapolitics news	-3.093*** (0.938)	-5.336*** (1.239)
Constant	-0.613 (0.708)	0.931 (0.891)
R-squared	0.033	0.068
<i>Panel C. Exposure before election</i>		
Dummy news before elections	-3.630 (2.192)	-2.383** (0.930)
Constant	-4.454 (2.278)	-3.372** (1.438)
R-squared	0.100	0.189
Time fixed effects	Yes	Yes
Candidate controls	Yes	Yes
Observations (Panels A & B)	356	356
Observations (Panel C)	119	119
Number of candidates (Panels A & B)	254	254
Number of candidates (Panel C)	77	77

Notes: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Each dependent variable is the municipal characteristic in the column header weighted by the candidate's municipal vote share each electoral year (see equation 10). The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel A, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In Panel B, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. In Panel C, 'News before elections' is a dummy variable that equals 1 if the candidate was linked to paramilitaries before the corresponding election. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.

Table C-6: Verifying media bias: Parapoliticians' vote concentration using *El Espectador* – Candidate-level regressions

Dependent variable:	Shade		Sunlight	
	Paramilitary presence (1)	Dummy high paramilitary presence (2)	State capacity (3)	Dummy high state capacity (4)
<i>Panel A. Exposure based on El Espectador</i>				
Parapolitician Dummy	1.203*** (0.212)	11.07*** (1.800)	-10.96*** (1.327)	-12.93*** (1.866)
R-squared	0.102	0.135	0.206	0.220
<i>Panel B. Exposure based on El Tiempo</i>				
Parapolitician Dummy	1.198*** (0.296)	9.765*** (1.887)	-9.170*** (1.585)	-11.824*** (1.858)
R-squared	0.105	0.115	0.183	0.208
Observations	339	339	339	339
Number of candidates	244	244	244	244

Notes: Robust standard errors in parentheses. *** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level. Each dependent variable is the municipal (shade or sunlight) characteristic in the column header weighted by the candidate's municipal vote share each electoral year (see equation 10). The sample of politicians included are all candidates to the Senate in the 2002, 2006 and 2010 elections with previous experience either in the House or Senate. In Panel A, 'Parapolitician Dummy' equals 1 if there are any stories connecting the candidate to paramilitaries at any point in time. In Panel B, 'Percent parapolitics news' is the percentage of total news stories of the candidate that relate him with the parapolitics scandal. Candidate controls include: number of news stories about the politician, gender, profession, years in politics, periods in office, and incomplete term dummy.