

Dead Watchdogs Don't Bark: How declining local news harms local voters, national parties and decentralization

Jori Korpershoek

February 29, 2024

1 Introduction

Like most politicians who make their way into national politics, Barack Obama did not start his career in the national capital. Rather, he needed a springboard, a place to “pay [his] dues” before making his way to higher office. In his case, this springboard was the State Senate of Illinois, “notorious for old-style pork- barreling, logrolling, payola, and other political mischief” (Obama, 2020). For an ambitious politician like Obama, this had two implications. First, he had to make his mark as a legislator, which he did by, among other things, help pass healthcare and anti-corruption reform. Second, it meant that, given the reputation of the state, he would have to make sure to do this while staying away from any legislation or dealmaking that would make him look corrupt in the eyes of national party elites, donors and future constituents. What did this mean for his effectiveness as a lawmaker? To quote a fellow (Republican) state senator on Obama’s local record: “it’s not so much what he did, but you have to sort of look at what he didn’t do” (Arons et al., 2008). According to his own biography, there were areas of legislation where he actively avoided getting involved.¹ Why did he avoid wheeling and dealing in particular policy areas? Because, quoting the State Senate whip, “Barack’s different, he’s going places” (Obama, 2020). He sacrificed influence at the state level for a clean reputation and a shot at higher office in the future.

Decentralization, giving power to states, municipalities and other subnational governments, is often justified because politicians at these levels are “closer to the people” and thus have the local knowledge to tailor policies to local needs in a way the national government can not. The example of the Illinois Senate points to two ways in which this can break down without high quality information about the decisions of politicians. First, corruption and other forms of abuse of office are an obvious waste of public funds. Second, career incentives are distorted because good types forego local influence to signal their disinterest in corruption to national elites. Local politicians of all stripes then choose to not optimally tailor policy to local needs at all, either to line their own pockets or to avoid the suspicion of doing so.

Why is this important? First, because in most western democracies, municipalities, states, provinces and other subnational entities play an important (and in many nations growing) role in the policy making process. Often justified because these layers of government are ‘closer to the people,’ these layers of government play a large part in public goods provision, including education, housing and infrastructure (Martinez-Vázquez et al., 2017). Second, there is a

¹One area where he had to be careful was pork barrel spending: some of the spending he secured for his district did end up going to financial backers. During the 2008 campaign it became an issue whether this had ‘just’ been pork flowing back to his district, or whether these were kickbacks in exchange for financial support (Long, 2007) As for other policy areas he avoided, in his autobiography, he gives the relatively example of gambling reform. His colleague on the other hand argues that he “went along with the program” in most policy areas.

growing crisis in local journalism. In the United States alone, over 1.800 communities lost their newspaper between 2008 and 2020. Newsroom employment at newspapers declined 57%, with online publications only picking up some of the slack (see Walker (2021), Hamilton (2016)). Other countries have similar trends (Nielsen, 2015). This means that while local governments have become increasingly important in policy making, there are also fewer eyes watching their decisions. And this has real effects: a growing empirical literature shows that decline in local news increases both corruption and party line voting at the local level (see e.g. Gavazza et al. (2019), Gao et al. (2020)) while a political science literature blames the decline of local news for the nationalization of local politics (see e.g. Hopkins (2018)).

The aim of this paper is to examine the implications of the decline of local journalism on the quality of local policy making, career incentives and how this affects the rationale for decentralisation. It does this by developing a two-stage accountability model, where at the end of the first period a local politician is either up for re-election (electoral incentives) or for a promotion to higher office (career incentives). It then proceeds in three steps. First, it investigates how the decline of local news leads to worse local behavior, focussing on how it distorts the incentives for well meaning politicians. They give up on tailoring policy, instead following the national party line in order to maintain a clean reputation and secure a shot at promotion. Second, it argues that given these dynamics, local voters might prefer voting for local or outsider candidates who have no chance of being promoted, but also do not feel to cater to national elites. Finally, I place these findings inside a simple fiscal federalism model and argue that without strong local news, the rationale for decentralization is undermined for two reasons. One, because politicians either abuse local information or choose not to use it. Two, because voters purposefully shut down the pipeline between local and national politics.

2 Decentralization, local news and parties

Whether under the banner of Federalism in the United States, the subsidiarity principle in Europe or ‘community-based governance’ elsewhere, subnational governments play an (increasingly) important role in many nations (see e.g. Martinez-Vázquez et al. (2017)). Arguments for decentralization often, implicitly or explicitly, rely on asymmetric information. Certain policy decisions should be made at the local level because local politicians know more about local preferences and circumstance. They thus can deliver better policy outcomes relative to their colleagues in distant capitals who are limited in how much they can tailor policy to local needs (see e.g. Oates (1972); Alesina and Spolaore (2003)). Take for example public procurement for a local infrastructure project. A national government could hire a big company to build a local road, but a local politician might have a better understanding of local needs, can pick a better site for the project and use their local network to find the company best suited for the job. This information asymmetry makes it “very difficult for external observers (citizens or analysts) to tell the difference between productive [and] counterproductive spending.” For example, “transportation infrastructures can remedy market failures, but they can also be a vehicle for corruption or for providing benefits to a geographic constituency” (Tabellini, 2005). If decentralization is motivated by the ‘local knowledge’ of politicians, this facts points to the importance of local watchdogs, who similarly understand local circumstance and local networks and can in

their investigations distinguish between hiring the best local company from clientelism.

At the same time, there is a growing crisis in local journalism, one of the leading contenders of checking and investigating local government. Local news production is a public good, long funded by the local information and advertising monopoly held by local newspapers. With increased competition, fueled by among many other things, the rise of the internet, profit margins are eroding and less money is going to local investigative journalism (Cagé, 2020; Angelucci and Cagé, 2019; Nielsen, 2015). In an evocative restatement of the losses in the newspaper industry, Doctor (2015) estimates that the lay-offs in the United States have led to a “cumulative loss of about 216,000 years of local knowledge.” Although some innovation in business models has occurred, this decline is visible in almost every western democracy.²

There is an extensive empirical literature that investigates the impact of changes in the media landscape on local politics. All of it finds that a decrease in the quality of local news leads to worse local policy making, but points to different channels. The first channel is obvious and is theoretically well understood: as corruption goes undiscovered, it becomes more likely (see e.g. Besley (2006); Drago et al. (2014); Gao et al. (2020)). This points to an increase in abuse of office. The second channel finds that a decrease in local news leads to an underuse of local office. For example, Gavazza et al. (2019) finds that the crowding out of local newspapers (by the roll out of broadband internet) led to a causal decline in the size of local government. Even though the *de jure* responsibility of localities did not change, spending on housing, social services and to a lesser extent, education, decreased. Moskowitz (2021) find that a decline in local reporters leads to less coverage of corruption and more party-line voting by politicians. These papers argue that as voters have less information due to the decline of local news, their representatives are more likely to get “in formation” and base their decisions on the party’s, rather than their constituents, needs (Trussler, 2020). This implies that the decline of local newspapers does not just lead to abuse, but also an *underuse* of office. Politicians ignore local circumstance and politics loses its local nature.

The empirical and theoretical literature on impact of news on accountability often often focus on how it effects electoral incentives(see e.g. Moskowitz (2021) and Hopkins (2018)). This ignores another important source of accountability: political parties and the career incentives they offer. The federalism literature has long stressed the importance of political parties as a second principal to local politicians. Going back to Riker (1964), scholars have argued that “the distribution of power in political parties” is an important determinant of the effectiveness of decentralization (Riker, 1975). *De jure*, local governments can have a lot of power, but if the parties within which politicians operate incentivize them to ignore local information, *de facto* power remains at the national level.³ How can parties exert influence? In most democracies, a large majority of politicians elected to national office have their start at a lower level of government.⁴ Impressing the selection committee of the national party, national donors and

²It is not obvious that online journalism can build up the same local social capital. As Simon (2009) noted in a Senate Hearing; “you do not – in my city — run into bloggers or so-called citizen journalists at City Hall, or in the courthouse hallways or at the bars and union halls where police officers gather. You do not see them consistently nurturing and then pressing sources. You do not see them holding institutions accountable on a daily basis.”

³In an extreme example, the Soviet Union was, at least nominally, a federal union. But the Communist party ensured that this power would not be used.

⁴In Sweden and the U.S., this is over 70% (Dal Bó et al., 2017; Borchert, 2009). In Germany and Brazil, it is

other powerbrokers will then be part of the calculus of the politician early in their career. The hope is the career incentives parties offer align the “political incentives of local politicians with national objectives” (Enikolopov and Zhuravskaya, 2007). However, what shapes these career incentives, which direction it pushes politicians in, and which politicians respond is often less clear.

This is an even more important question because even though the academic literature often points to career incentives as beneficial, a lot of anecdotal evidence also points to its limitations. One of the main reasons career incentives are seen as useful is because higher office is often seen as a bigger prize than local office, so they give an additional reason for politicians in local office to behave well. For example, Myerson (2006), develops a theoretical model in which career incentives are vital to make both local and national politics work in democracies. However, he does so in a setting without uncertainty about the actions of politicians: good and bad policy decisions can easily be distinguished. But as highlighted above, this asymmetric information about local circumstance is one of the motivating reasons for decentralization. What happens if parties do not have this information? A rising critique of political parties is that they are “too centralized, failing to empower [local chapters] and party members,” while career politicians are seen as catering more to the party than their (local) constituents (Richards, 2017). Two quips capture the sentiment. On the one hand, the poor performance of a Baltimore mayor was blamed on his ambition: “what happened under his watch (...) was that he wanted to be governor” (Keller, 2015). On the other, the performance of Michael Bloomberg in New York was seen as standing in the way of his future political career: “what makes him good at being mayor hinders his presidential ambitions” (Barber, 2013). The success of local parties as well as independent and outsider candidates at the local level point to a revealed preference for politicians without any career incentives, at least in some circumstances.

This paper argues that the gap between these arguments can be bridged by considering the information national political parties have. They too are “external observers:” they too have limited ability to discern the difference between optimal adaptation to local circumstance and abuse of office (see e.g. Fowler and McClure (1989), chapter 7).⁵ As Lorentzen (2014) argues in the case of China, local news can be a vital source for national actors about the actions of local politicians. I argue that the less information national parties have about local politics, the less they trust the use of local information in local policy making and the less willing they are to promote these candidates. Politicians interested in higher office then start to follow the national party line in order to avoid suspicion. The reason career incentives are costly for voters is not just that they push politicians to ignore local information. It is which politicians choose to respond. Since national parties and local voters are ultimately interested in selecting the same type of politician (for example, both dislike corrupt politicians), the ‘distortion’ is at the top: it is exactly politicians like Obama, talented and interested in good policy, who give up their local influence in order to impress national elites and have a shot at making a bigger impact later in their career. The decline of local news opens up a “void” between national parties (and their

over 60% (Borchert, 2009).

⁵This chapter describes in the lead up to a primary for a Congressional seat in 1984 and 1986, national donors were more likely to give support to candidates who looked good on paper, rather than those who fit the district. Choices that made sense given local circumstance but were out of step with the national environment (e.g. switching parties, being soft on certain kinds of crime) were deemed to make a candidate unelectable.

local candidates) and the local electorate (see e.g. Mair (2013)). In extremis, voters might even turn away from established parties and elect candidates who face no career incentives at all.

3 Model set-up

The political economy is made up of a single nation with many regions. I focus on a generic region where a politician is in charge of providing a continuum of local public goods $x_{i,t}$, $i \in [0, 1]$, in each period $t \in \{1, 2\}$. Think of these as the many decisions any local politician has to make during their term, including: what types of schools to open and where to build them, when and how the trash is be collected, and deciding which companies should be hired to build local infrastructure.

There is an incumbent in office during the first period. At the end of the first period, there is an election. With some probability (defined below) the national party needs a candidate for national office and considers whether or not to promote the politician. If the politician gets the candidacy, they can not run for re-election.

There is a local representative voter, who cares about local policy outcomes ($w(x_{i,t})$).

$$V = \int_0^1 w(x_{i,1})di + \delta \int_0^1 w(x_{i,2})di$$

These policy outcomes depend on the policy choices of the politician elected in period 1 and period 2.

While voters only care about policy outcomes, politicians have different interests. There are two types of politicians, good and bad ($T \in \{G, B\}$). The probability that any politician is the good type is π . Good politicians care about policy outcomes while they are in office ($w(x_{i,t})$), the value of being in local office (v_L) and the value of running for national office (v_N). Bad politicians care about personal rents ($r(x_{i,t})$), the value of being in local office and the value of running for national office. Their respective utility functions are are:

$$U_G = v_L + \int_0^1 w(x_{i,1})di + \delta \left[\left(v_L + \int_0^1 w(x_{i,2})di \right) \cdot \mathbf{1}_R + v_N \cdot \mathbf{1}_N \right]$$

$$U_B = v_L + \int_0^1 r(x_{i,1})di + \delta \left[\left(v_L + \int_0^1 r(x_{i,2})di \right) \cdot \mathbf{1}_R + v_N \cdot \mathbf{1}_N \right]$$

Where $\mathbf{1}_R$ is an indicator function that is 1 if the politician is re-elected in the second period and 0 otherwise, and $\mathbf{1}_N$ is an indicator function that is 1 if politician is nominated for national office and 0 if they are not. I assume that $v_N > v_L$: politicians prefer holding national over local office.

The political party on the other hand only cares about local office as a source for candidates for higher office. With probability ρ national parties they need such a candidate. They want to nominate good politicians. Conditional on needing a candidate, their utility is:

$$P(T) = \begin{cases} 1 & \text{If they promote a politician of type } G. \\ 0 & \text{If they promote a politician of type } B. \end{cases}$$

Type (x_i)	Policy pay-off ($w(x_i)$)	Rent ($r(x_i)$)	Verifiability
Good policy (g)	g	0	Only verifiable by the newspaper.
Bad policy (b)	0	r	Only verifiable by the newspaper.
Uniform policy (u)	$u(< g)$	0	Immediately verifiable.

Table 1: The three policies between which local politicians can choose.

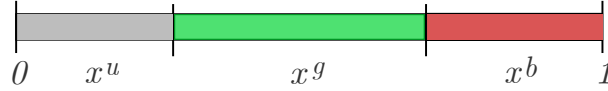


Figure 1: The local politician controls a range of public goods ($i \in [0, 1]$). For each of these public goods they can implement it as a good policy (g), a bad policy (b) or a uniform policy (u). Their decision can be summarized as the ranges, x_u , x_g and x_b respectively, for which they choose to implement these different policies.

If they do not reelect the current incumbent, their candidate is a draw from the pool of potential politicians.

The role of the politician is to implement local public goods. For each of the public goods in the range $[0, 1]$, the politician can implement the public good $x_{i,t}$ in three different ways: $x_{i,t} \in \{g, b, u\}$. For example, say that one of the public goods they need to provide is a road and they need to choose a contractor to build it. Then they can choose a:

- Good policy (g). The politician uses their local network to find the most efficient contractor with a lot of local experience. This delivers policy pay-offs $w(g) = g$ and no personal rent to the politician $r(g) = 0$.
- Bad policy (b). The politician hires a crony from their local network in exchange for a kickback. This generates no policy pay-offs, $w(b) = 0$, but gives personal rent $r(b) = r$ to the politician.
- Uniform policy (u). The politician hires a large national company without local experience. This generates some policy pay-offs $w(u) = u$ and no personal rent.

I refer to the first two as bottom-up public goods provision, since they rely on the local network of the politician. The uniform policy I call top-down or following the national party line, since it does not rely on local knowledge. The rationale for decentralization of local public goods provision in this model is that local politicians can make better choices. Hiring the national company could have simply been done by a national politician without local knowledge. So I assume that the politician using their local connections is optimal: $g > u$. However, when the politician uses their own network, citizens and parties do not know whether they used it to find the most efficient contractor or to enrich themselves unless the newspaper reports on this. If the politician chooses the uniform policy on the other hand, this immediately becomes common knowledge. The policy types, their pay-offs and verifiability are summarized in Table 1.

Politicians choose how to implement public goods. What matters for pay-offs is the ranges of public goods ($\in [0, 1]$) for which they choose the uniform, the good and the bad policy. Denote these ranges as $\{x^u, x^g, x^b\}$, with $x^u + x^g + x^b = 1$. See Figure 1.

Before the election and the promotion decision, the newspaper investigates the choices of the politician in the first period. For all policies where the politician chose the uniform policy ($x_{i,1} = u$) the newspaper cannot report anything new. For all remaining choices, the newspaper investigates implemented the good policy ($x_{i,1} = g$) or the bad policy ($x_{i,1} = b$). With probability $\gamma \in [0, 1]$ all remaining policy choices become common knowledge. γ is interpreted as the quality of local news.⁶

Timing

Period 1

1. Nature draws the type of the politician ($T \in \{G, B\}$). Politicians learn their type.
2. For each public good under the local politician's control ($x_{i,1}, i \in [0, 1]$) the politician chooses how to implement it ($x_{i,1} \in \{g, b, u\}$). For all public goods where the politician chooses u , this becomes common knowledge.
3. The newspaper investigates the remaining choices of the politician. With probability γ the remaining choices of the politician become common knowledge.
4. With a probability ρ a national party need a candidate for higher office. They update their beliefs about the type of politician using Bayes' rule using the politician's policy choices for u and the newspaper report. They choose between offering the candidacy to the incumbent and taking a random draw from the pool of potential politicians.

Whether or not there was a promotion opportunity is common knowledge.

5. There is an election. If the politician is not promoted, voters use Bayes' rule to update their beliefs about the type of the politician using the newspaper report, the promotion decision of the party and the policy choices for u . They choose between re-electing the incumbent or voting for a challenger.

If the incumbent is promoted, they elect a new politician, who is a random draw from the pool of potential politicians.

Period 2

1. If a challenger wins the election, they learn their type. Whoever is in office chooses the policies that maximize their utility.
2. Outcomes materialize.

3.1 Equilibrium concept

I solve the game for all Perfect Bayesian Nash Equilibria, restricting myself to those that follow a 'natural language.' This rules out equilibria in which good politicians choose bad policies just because it is expected of them. Equilibria in this game then come in three varieties. Pooling equilibria, Standard Separating equilibria and what I will call Party Discipline equilibria.

⁶A natural alternative would be to assume that the probability of successfully uncovering the behavior of the politician depends on their behavior. For example, that if the politician was corrupt for a share of policies x^b , the probability of discovery could be $\gamma \cdot x^b$: The more corruption, the easier it is to discover. This would strengthen the results, but comes at the cost of tractability.

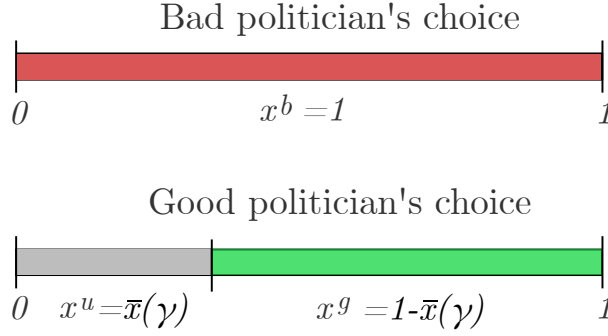


Figure 2: The first period policy choice. The bad politician always chooses the corrupt policy. The good politician chooses the uniform policy for a range \bar{x} of policies such that the bad politician does not want to mimic, and the good policy for all remaining public goods.

The Pooling and the Standard Separating equilibrium are the usual varieties found in the literature, like Besley (2006). They are fully derived in Appendix A. In a Pooling equilibrium, both good and bad politicians choose the same policies. This means no information about the type of the politician is ever revealed and both good and bad types are re-elected or promoted. In a Standard Separating equilibrium, good politicians choose good policies, bad politicians choose bad policies and information is revealed only if the newspaper reports on the choice ⁷

The addition of the uniform policy, which is immediately verifiable, adds a third type of equilibrium to this game. I call this the Party Discipline equilibrium. In it, good politicians signal their type by following the national party line and choosing the uniform policy. Bad types choose not to follow the party line and reveal their type in the process. It is thus also a separating equilibrium, but, because the uniform policy is immediately verifiable, one in which the type of the politician is revealed independent of the newsreport. I look for the conditions under which this separating equilibrium exists, in which:

- Politicians choose the policies that maximize their pay-offs for all public goods $x_{i,t}$, $i \in [0, 1]$, $t \in \{1, 2\}$. Then in a separating equilibrium:
 - In both period 1 and 2, bad politicians choose the bad policy for all policies ($\{x^u = 0, x^g = 0, x^b = 1\}$), given their re-election incentives, career incentives and the beliefs of parties and voters.
 - In period 1, good politicians choose the uniform policy for a range $x^u = \bar{x}(\gamma, r, v_L, v_N) \in [0, 1]$ and the good policy for the remainder ($\{x^u = \bar{x}, x^g = 1 - \bar{x}, x^b = 0\}$), given their re-election incentives, career incentives and the beliefs of parties and voters. In period 2, they choose the good policy for all public goods.
- Denote as $\hat{\pi}_P$ the belief that the party has that the politician is the good type. Parties believe the politician is the good type ($\hat{\pi}_P = 1$) and promote them if they observe the politician chose the uniform policy for a range of policies $x^u = \bar{x}(\gamma, r, v_L, v_N)$.
- Denote as $\hat{\pi}_V$ the belief that voters have that the politician is the good type. Voters believe the politician is the good type and re-elect them if the politician was not promoted and

⁷To be precise, in these equilibria politicians might choose some uniform policies, this choice is just not informative.

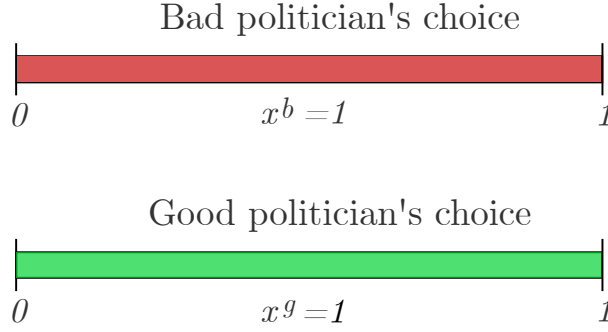


Figure 3: The second period policy choice. The bad politician chooses the bad policy for all public goods. The good politician chooses the good policy.

the politician chose u for a range of policies $x^u = \bar{x}(\gamma, r, v_L, v_N)$.

4 Solution: Topdown separating equilibrium

I solve the game by backwards induction where possible. In this equilibrium good types signal their type by choosing the verifiable uniform policy and foregoing some policy rents in period 1, in order to secure reelection or promotion for the next period. This can be a credible signal, since good types give up only some policy pay-offs (they get u for these policies instead of g) while bad types would have to give up all corruption rents (getting 0 instead of r).

Period 2

In the second period, politicians in local office can no longer be re-elected or promoted. They simply choose to implement the policies that maximize policy pay-offs (if they are the good type) or corruption rents (if they are the bad type). The problem of the good politician is then:

$$\max_{x_{i,2} \in \{g, b, u\} \forall i \in [0,1]} v_L + \int_0^1 w(x_{i,2}) di$$

Given that pay-offs are linear and good policy delivers the highest policy pay-offs (since by assumption it is better than the uniform policy, $g > u$, and the bad policy has no policy pay-offs), this is equivalent to the following:

$$\max_{x^u, x^g, x^b} v_L + x^u \cdot 0 + x^b \cdot 0 + x^g \cdot g \quad \text{s.t. } x^u + x^g + x^b = 1$$

From which it is clear that the politician chooses the good policy for all public goods. Their optimal second period strategy is thus $\{x^u = 0, x^g = 1, x^b = 0\}$.

Bad politicians on the other hand maximize the amount of corruption rents they get.

$$\max_{x^u, x^g, x^b} v_L + x^u \cdot 0 + x^b \cdot r + x^g \cdot 0 \quad \text{s.t. } x^u + x^g + x^b = 1$$

From this it follows that in period 2, bad politicians bad policy for all public goods, since that is the only policy that generates personal rent. The strategy they choose is $\{x^u = 0, x^g = 0, x^b =$

1}

Period 1

I now will show that good politicians can effectively signal their type by following the party line. I will do this by first proposing beliefs and deriving optimal behavior of parties and voters, taken as given these beliefs and that good politicians follow the party line for a share of policies $\bar{x}(\gamma)$, while bad politicians do not follow the party line at all. Then I show under which conditions it is optimal for politicians to behave this way and derive the minimal level of party line voting ($\bar{x}(\gamma)$) consistent with this separating equilibrium, given the beliefs and strategies of voters and parties.

Period 1: Election

Because voters care about policy pay-offs, they want to elect a politician who will choose good policies in period 2. Since the good politician chooses good policies in period 2, while the bad politician chooses bad policies, this amounts to trying to elect good politicians.

How do they determine what the type of the politician is? When the election happens, they have three pieces of information. First, they can observe whether or not the politician was promoted. Second, if there was a news report, they know whether the politician chose the good or the bad policy for all policies where the politician chose the bottom-up policy. Finally, for all policies where the politician chose the top-down option (the uniform policy), voters can immediately verify this choice.

In a Party Discipline equilibrium in which good politicians signal their type by choosing the immediately verifiable uniform policy u for a range of policies \bar{x} , it is this choice that reveals information. If good politicians follow the party line and bad politicians do not, once you observe a politician following the party line, you know they are the good type. This is independent of what the newspaper reports, since the uniform policy is immediately verifiable. The newspaper might uncover wrongdoing or the good choices of a good type, but this would only confirm what voters already know.

$$\hat{\pi}_V(x^u) = \begin{cases} 1 & \text{If } x^u \geq \bar{x} \text{ and the newspaper did not find abuse of office} \\ & \text{or if the newspaper confirms } x^g = 1. \\ 0 & \text{Otherwise.} \end{cases}$$

Voters thus re-elect politicians who chose $x^u \geq \bar{x}$. Of course, they can only re-elect politicians who was not promoted.

Period 1: Promotion

The logic of the choice of political parties whether or not to promote the politician is similar. They can rely on the newspaper report and what they learn from the choices of politicians. Again, in a Party Discipline equilibrium, all information is revealed through the choice. Parties believe the politician is the good type if they followed the party line on a sufficiently large

number of policies.

$$\hat{\pi}_P(x^u) = \begin{cases} 1 & \text{If } x^u \geq \bar{x} \text{ and the newspaper did not find abuse of office,} \\ & \text{or if the newspaper confirms } x^g = 1. \\ 0 & \text{Otherwise.} \end{cases}$$

Period 1: The policy choice

This leaves the choice of the politician. For every policy under their control, politicians choose either the good policy, the bad policy or the uniform policy. The key to the Party Discipline equilibrium is that good politicians can credibly signal their type to voters and the party by choosing the uniform policy. But how much do they have to follow the national party line to reveal their type?

Before deriving this, there are two important preliminary steps. First, given that they control a continuum of policies, the set of possible strategies is large. In Lemma 1 I show that the set of relevant strategies is a lot smaller. For example, in any equilibrium, the good politician only considers two strategies. Second, I derive parameter conditions in which this equilibrium exists.

Lemma 1. *In any equilibrium, the only relevant strategies that politicians consider are:*

- *Good politicians only consider to follow the party line for the minimal expected level and choose good policies for the rest ($\{x^u = \bar{x}, x^g = 1 - \bar{x}, x^b = 0\}$) or to choose the good policy for all policies ($\{x^u = 0, x^g = 1, x^b = 0\}$).*
- *Bad politicians only consider to be fully corrupt ($\{x^u = 0, x^g = 0, x^b = 1\}$), to partially mimic good politicians and be corrupt for the remaining policies ($\{x^u = \bar{x}, x^g = 0, x^b = 1 - \bar{x}\}$) or to pick a strategy that is pay-off equivalent to fully mimicking the good politician ($\{x^u = \bar{x}, x^g = 1 - \bar{x}, x^b = 0\}$).*

Proof. See Appendix. □

The intuition behind this lemma is simple: since both policy pay-offs and corruption rents are linear in the choice of the politician, both the good and the bad politician only consider the highest possible level of x^g and x^b , given equilibrium beliefs, career incentives and electoral incentives. On the other hand, if the bad politician does not choose b at all, all choices that meet the constraint that $x^u \geq \bar{x}$ are pay-off equivalent, since neither good nor uniform policies generate corruption rents.

For following the party line to be a credible signal of type, bad politicians should not always want to mimic good politicians. The following parameter restriction is required for the existence of a separating equilibrium in which the good politician uses the uniform policy to signal. It imposes that corruption rents are sufficiently high: without this, bad politicians would always mimic good politicians and there is no separating equilibrium where u is chosen only by good politicians. In other words, the parameter restriction rules out pooling equilibria.

Lemma 2. *There is only a top-down separating equilibrium if corruption rents are sufficiently high $r \geq \frac{\delta (\rho v_N + (1-\rho) v_L)}{1-\delta (1-\rho)}$*

Proof. For the existence of a Party Discipline equilibrium, it is a necessary condition that fully mimicking the good type and foregoing all corruption rents in period 1 is a dominated strategy. This strategy induces positive beliefs, and there is no risk of getting caught by the newspaper. This strategy is dominated for bad politicians if:

$$\begin{aligned} U_B(x^u = 0, x^g = 0, x^b = 1) &\geq U_B(x^u = \bar{x}, x^g = 1 - \bar{x}, x^b = 0) \\ r + v_L &\geq 0 \cdot r + v_L + \delta(\rho v_N + (1 - \rho)(r + v_L)) \\ r &\geq \frac{\delta(\rho v_N + (1 - \rho)v_L)}{1 - \delta(1 - \rho)} \end{aligned}$$

Note that this also immediately rules out any strategy $\{x^u = \bar{x} + \alpha, x^g = 1 - (\bar{x} + \alpha), x^b = 0\}$, $\alpha \in [0, 1 - \bar{x}]$, since these are all pay-off equivalent. \square

Assumption 1. *Corruption rents are sufficiently high.* $r \geq \frac{\delta(\rho v_N + (1 - \rho)v_L)}{1 - \delta(1 - \rho)}$

Suppose the good politician tries to signal their type by following the party line on a range \bar{x} of policies. Then if Assumption 1 holds, then the bad politician has two relevant strategies. In the first, they follow the party line to the same extent as the good politician ($x^u = \bar{x}$). But for all remaining public good they choose the bad policy. Then as long as the newspaper does not uncover their wrongdoing, they are thought of as good politicians. Second, they can choose to not follow the party line at all and choose the bad policy for all public goods. By not following the party line they always reveal their type, but maximize their first period pay-offs.

Suppose that parties and voters expect politicians who choose the uniform policy for a range \bar{x} to be good types. How high should \bar{x} for bad politicians not to want to mimic and induce these positive beliefs?

Proposition 1. *The minimal level of top down policy making consistent with a Party Discipline equilibrium is $\bar{x}(\gamma) = \frac{1}{r}(1 - \gamma)\delta[\rho v_N + (1 - \rho)(r + v_L)]$.*

Proof. In this equilibrium, bad politicians choose to be fully corrupt, choosing the bad policy for all policies ($\{x^u = 0, x^g = 0, x^b = 1\}$). Independent of what the newspaper reports, the choice to not choose the top-down policy reveals they are the bad type. This means they never get reelected or promoted. Their alternative is to partially imitate good types: choose the uniform policy for some public goods, the corrupt policy for the rest and hope the newspaper does not catch them so they secure reelection or a promotion ($\{x^u = \bar{x}, x^g = 0, x^b = 1 - \bar{x}\}$). They do not do this if:

$$\begin{aligned} U_B(x^u = 0, x^g = 0, x^b = 1) &\geq U_B(x^u = \bar{x}, x^g = 0, x^b = 1 - \bar{x}) \\ r + v_L &\geq (1 - \bar{x})r + \bar{x} \cdot 0 + v_L + \delta(1 - \gamma)(\rho v_N + (1 - \rho)(r + v_L)) \end{aligned}$$

Solving this for \bar{x} tells us how high \bar{x} has to be for top-down policy making to be a credible signal of type:

$$\bar{x}(\gamma) = (1 - \gamma) \frac{\delta}{r} (\rho v_N + (1 - \rho)(r + v_L))$$

\square

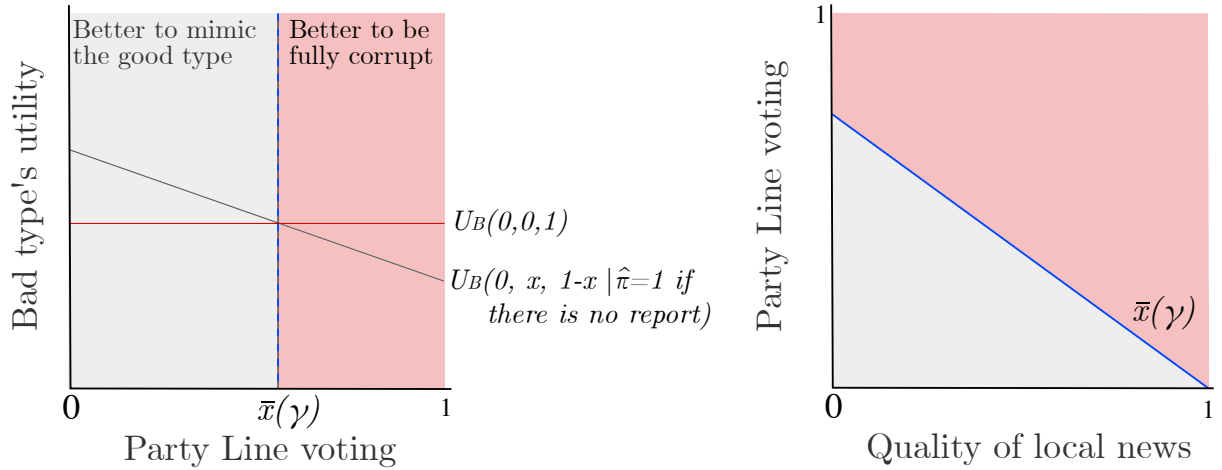


Figure 4: The utility from being fully corrupt or mimicking the good type as a function of the amount of top-down voting expected. The more top-down voting is expected, the less attractive it is to mimic the good politician.

This is the minimum amount of uniform policy making that good types have to follow to signal their type (see Figure 4). Any lower and bad types would mimic their behavior, making it an ineffective signal.

Lemma 3. *The level of top-down policy making required for the good politician to signal their type is decreasing in the quality of local news (γ) and corruption rents (r), and increasing in the value of local and national office.*

Proof.

$$\begin{aligned} \frac{d\bar{x}(\gamma)}{d\gamma} &= -\frac{\delta}{r}(\rho v_N + (1-\rho)(r + v_L)) < 0 & \frac{d\bar{x}(\gamma)}{dr} &= -(1-\gamma)\frac{\delta}{r^2}(\rho v_N + (1-\rho)v_L) < 0 \\ \frac{d\bar{x}(\gamma)}{dv_N} &= (1-\gamma)\frac{\delta}{r} \cdot \rho > 0 & \frac{d\bar{x}(\gamma)}{dv_L} &= (1-\gamma)\frac{\delta}{r} \cdot (1-\rho) > 0 \end{aligned}$$

□

In equilibrium, the main alternative strategy for bad politicians is to follow the party line as much necessary and to be corrupt for all remaining policies, hoping to not be discovered by the newspaper. The higher the probability of discovery, the less attractive this deviation is. This means that if the quality of local news goes up, the amount of party line voting needed decreases. The value of local and national office increase the level of party line voting required, because it increases the benefits of holding office in period 2. The value of corruption rents reduce the level of party line voting necessary, because it makes *not* following the party line and maximizing first period rents more attractive.

The question then becomes when good politicians are willing to choose the uniform policy to signal their type. It is costly for them: By choosing u rather than g , they get lower policy pay-offs in period 1. These costs depend on the quality of local news: the lower the quality of local news is, the higher these costs are. At the same time, there is the benefit of signaling that they

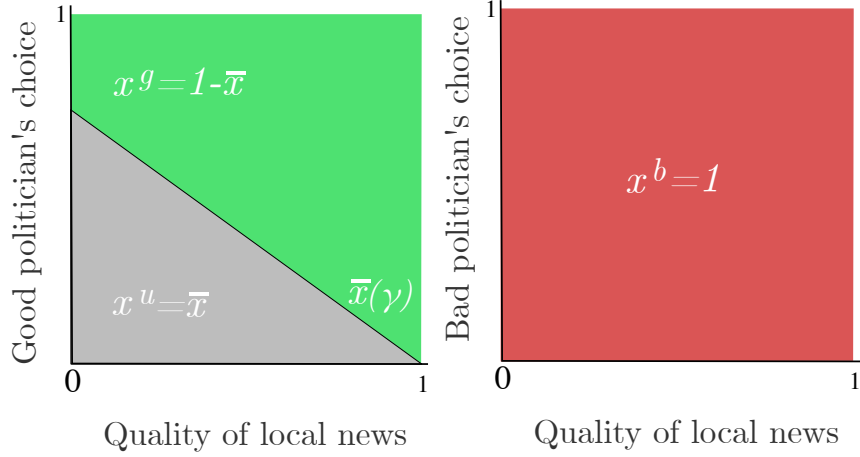


Figure 5: The first period choices of politicians as a function of the quality of local news. The lower the quality of the newspaper, the more good politicians have to choose the uniform policy to signal their type.

are a good type, which brings certainty that they will be reelected, or promoted if the chance arises. But if the costs outweigh these benefits, good politicians would rather maximize their first period impact by only choosing good policies, and giving up on either getting reelected or promoted unless the newspaper reports on their good behavior. The following condition pins down when good politicians are willing to follow the party line:

$$\begin{aligned}
 U_G(\{x^u = \bar{x}, x^g = 1 - \bar{x}, x^b = 0\}) &\geq U_G(\{x^u = 0, x^g = 1, x^b = 0\}) \\
 \bar{x} u + (1 - \bar{x})g + v_L + \delta[\rho v_N + (1 - \rho)(v_L + g)] &\geq g + v_L + \gamma \delta(\rho v_N + (1 - \rho)(v_L + g)) \\
 g &\leq u + r \cdot \left(1 + \frac{(1-\rho) u}{\rho v_N + (1-\rho)v_L}\right)
 \end{aligned}$$

Assumption 2. *The benefits of choosing the good policy are not too high. $g \leq u + r \cdot \left(1 + \frac{(1-\rho) u}{\rho v_N + (1-\rho)v_L}\right)$*

If Assumption 1 and 2 are met, then there is an a Party Discipline equilibrium in which good politicians use the uniform policy to signal their type. Equilibrium behavior of politicians is summarized in Figure 5, for all possible values of the quality of local news. Bad politicians always abuse local office, independent of the quality of local news. This means that any politicians who engages only in bottom-down policy making is suspicious. Good politicians set themselves apart from bad politicians by choosing the immediately verifiable uniform policy u for a share of policies that bad politicians would not mimic. The lower the quality of local news, the larger this share has to be.

Proposition 2. *If Assumption 1 and 2 are met, there is a Party Discipline equilibrium in which good politicians use the uniform policy to signal their type by choosing policies $\{x^u = \bar{x}, x^g = 1 - \bar{x}, x^b = 0\}$. Both voters and parties believe they are good types if they choose this strategy or choose $x^g = 1$, and bad types otherwise. They are promoted with probability ρ and re-elected with probability $1 - \rho$.*

Bad politicians are fully corrupt in the first period and choose $\{x^u = 0, x^g = 0, x^b = 1\}$. They are never re-elected or promoted.

Proof. It follows from the previous propositions and lemmas. \square

4.1 Uniqueness and equilibrium refinement

Without further assumptions the equilibrium described is not unique. All equilibria of the game are described in Appendix A. The most plausible equilibrium to the one described above is a Standard Separating equilibrium, in which good politicians only choose good policies, bad politicians only choose bad policies and the newspaper is informative.⁸

For further refinement, I assume that the second period pay-offs of the politician depend on the beliefs that parties have about them. This is the case if the amount of effort the national party puts into advancing the career of a politician is higher if they are more confident they are the good type. Examples of this would be placing candidates higher on the national party list, nominating them for more valuable offices, or giving them more resources to campaign. If this is the case, and if this increase in pay-off is sufficiently valuable, the Party Discipline equilibrium is the only equilibrium that survives refinement by the intuitive criterion.

Assumption 3. *The value of running for national office is an increasing function of the belief of the party about the type of the politician. $v_N(\hat{\pi}_P = 1)$ is sufficiently larger than $v_N(\hat{\pi}_P = \pi)$:*

$$v_N(\hat{\pi}_P = 1) \geq \left(\frac{r}{r-1} \right) \left(v_N(\hat{\pi}_P = \pi) + \frac{1}{r\rho}(1-\rho)(v_L + r) \right)$$

Proposition 3. *If Assumption 1, 2 and 3 hold, the equilibrium described in Proposition 2, is the only equilibrium that meets the Intuitive Criterion (Cho and Kreps, 1987).*

Proof. See Appendix. \square

The equilibrium described above is the equilibrium that generates most information about the type of the politician at the lowest cost in terms of policy adaptation. If being promoted with positive beliefs is sufficiently valuable, good politicians want to deviate to this equilibrium, while bad politicians would not mimic their behavior (since \bar{x} is set such that they would not).

The difference between the Standard Separating equilibrium and the Party Discipline one are described in more detail in Section 5.2, where I give voters the opportunity to choose between electing a politician with and without career incentives.

4.2 Discussion

I will now discuss welfare for local voters, national parties and politicians in the Party Discipline equilibrium.

In equilibrium, local voter welfare is as follows.

$$V(\gamma) = \pi \left[(\bar{x}(\gamma)(g - u) + u) + \delta \cdot g \cdot ((1 - \rho) + \pi\rho + (1 - \pi)) \right] \quad (1)$$

⁸For the parameter values for which the Party Discipline equilibria exists rule out all pooling equilibria. The only other equilibria are separating equilibria in which all politicians follow the party line to some extent ($\bar{x}' < \bar{x}$) because it is expected of them, but where this does not contain any information.

Proposition 4. *Voter welfare is increasing in the quality of local news (γ), and decreasing in the value of a promotion (v_N), and increasing in the pay-offs of corruption (r).*

Proof.

$$\begin{aligned}\frac{dW}{d\gamma} &= \pi \cdot \frac{d\bar{x}(\gamma)}{d\gamma}(g - u) > 0 \\ \frac{dW}{dv_N} &= \pi \cdot \frac{d\bar{x}(\gamma)}{dv_N}(g - u) < 0 \\ \frac{dW}{dr} &= \pi \cdot \frac{d\bar{x}(\gamma)}{dr}(g - u) > 0\end{aligned}$$

All the signs follow from the signs of the derivatives derived in Lemma 3. □

If the quality of local news declines, voter welfare declines as well. Although the sign of this effect is the same as in canonical models (e.g. Besley (2006)), the mechanism is completely different. Here it is not that a decline in newspaper quality leads to less information being produced, but rather that the cost of producing this information goes up. So rather than a decline in local news leading to more corruption in period 2, the cost is limited to increased party line voting in period 1. The lower the quality of local news, the more suspicious parties (and voters) are of local policy adaptation, so the more good politicians have to follow the national party line to set themselves apart. A decline in local news thus leads to a ‘nationalization’ of local politics (see e.g. Hopkins (2018)).

Second, it is important to note that while the choice of the politician produces information about their type, this information most beneficial to voters if there is no promotion. If the good politician is promoted, local voters have no choice but to elect a new politician. An increase in the probability that the politician gets promoted thus hurts voters in period 2. Of course, local voters might internalize some of the benefits of the promotion of good politicians to national politics. But by construction, the benefits of national selection are shared among all regions, while local voters only internalize the local benefits.

For political parties on the other hand, welfare does not depend on anything apart from the probability that the politician is a the good type. This is because the equilibrium described is a separating equilibrium with full information, and for political parties only information matters.

$$P(\pi) = \pi(2 - \pi) \qquad \text{If there is a promotion}$$

This reveals a difference between voters and national parties. Voters care about local welfare, for which learning the type is useful, but comes at the expense of distorted choices in period 1. National parties on the other hand, only care about information, they have no ‘skin in the game’ when it comes to local voter policy decisions. Distortions that are costly for voters thus leave parties unaffected. This contrast is further discussion below, where I investigate whether voters would rather elect a politician from a national party, knowing that good politicians will distort their behavior in order to secure a promotion, or a politician from a local party, whose behavior is not distorted by career incentives.

Finally, it is useful to briefly discuss politicians’ equilibrium pay-offs. The utility of good

and bad types of politicians elected in period 1 is equal to:

$$U_B = v_L + r$$

$$U_G = \bar{x}(\gamma) \cdot u + (1 - \bar{x}(\gamma)) \cdot g + \delta(\rho v_N + (1 - \rho)(v_L + g))$$

Equilibrium pay-offs of bad politicians are relatively straightforward. Because their type is always revealed in equilibrium, they never get reelected or promoted. Their pay-offs thus depend only on the value of holding local office and the corruption rents they get when they are in office. For good politicians, the most interesting comparative static is that their pay-off from holding office decreases if the quality of local news declines. While good politicians care about the value of holding local office and future career benefits, they also care about policy outcomes. The more they have to distort their behavior, the worse off they are. Below, I will argue that this makes it less attractive for good types and might make it less likely that they enter politics in the first place.

5 The national implications of the decline of local news

In the following sections, I explore the national implications of the decline of local news. I do this in three steps.

First, I argue that a decline in local news worsens a selection at the local level. Specifically, since good politicians have to constrain their behavior in order to get elected, good politicians of high ability are less willing to run for local office. This has implications for local voters, but also impacts national parties, who lose out on running these talented politicians for higher office.

Second, given the downsides of career incentives discussed in the previous section, I explore the implications for competitions between different kinds of parties. Specifically, I give voters the choice between electing a politician from a national party, who faces career incentives and might get promoted, and a local party, who only faces electoral incentives. I show that as the quality of local news declines and the cost of career incentives increase, voters might prefer electing a local party or independent candidate.

Finally, I place these findings inside a model of fiscal federalism and decentralization, and describe how the decline of local news affects the standard rationale for decentralization.

5.1 Local news and selection into local politics

In many countries, local politics functions as a talent pool for the national arena. Whether or not well meaning talented people want to enter local politics will depend both on their long run career prospects (is there a chance to use local office as a springboard for national politics), but also on how rewarding it is to hold local office. The more local politicians have to follow the national party line in order to get reelected or promoted, the less attractive holding local office is. In this way, a decline in local news does not just affect the behavior of politicians in office, but might also affect who runs for office in the first place.

In Appendix B, I propose an extension which models the entry decision of politicians. I do this by adding a step prior to the game propose above, in which potential politicians decide

whether or not they want to run for office in the first place. The key insights from this extension come from the first period entry decision. It is only this decision that depends on the quality of local news. In the following paragraphs, I explain the changes to the pay-offs of the players and the implications for selection in period 1, while leaving period 2 selection, the promotion and reelection decision and the associated beliefs of parties and voters to the Appendix.

The main change in this extension is in the utility functions of politicians, voters and parties.

First, on the side of politicians. Let politicians differ in their ability, $a_j \sim U[1, 2]$. Ability does two things. First, it increases the effectiveness of their policy making, with policy made by higher ability politicians leading to higher policy and corruption pay-offs: $a_j \cdot w(x_i)$. Second, their ability influences their market salary, $a_j \cdot s$. Finally, I let the pay-offs of voters and parties directly depend on the ability of politician. For voters', it multiplies their policy pay-offs:

$$V = a_1 \int_0^1 w(x_{i,1}) di + \delta a_2 \int_0^1 w(x_{i,2}) di$$

Where a_1 is the ability of the politician in office in Period 1 and a_2 is the ability of the politician in office in Period 2. For parties it multiplies their pay-off of promoting a good politician:

$$U_N(T, a) = \begin{cases} a & \text{If they promote a politician of type } G \text{ and ability } a. \\ 0 & \text{If they promote a politician of type } B. \end{cases}$$

In the Appendix I show that from the perspective of voters' there is always positive selection in period 1, meaning that the pool of politician's running for office in period 1 is of higher ability than those who run in the second period, so that, given the chance, they would reelect the incumbent if they are revealed to be the good type. For the party, this is true as long as good types are relatively rare (so if π is relatively low). In that case, the the information revealed by being in office in period 1 is useful for parties.

If parties are indeed willing to promote the politicians who run in period 1, it is possible to pin down which politicians would run for office in period 1. Politicians are willing to run for office if the utility of being in office is higher than their outside option.

$$\begin{aligned} U_G(a_j, \gamma) &= a_j(\bar{x}(\gamma) \cdot u) + (1 - \bar{x}(\gamma)) \cdot g + v_L + \delta(\rho v_N + (1 - \rho)(v_L + a_j g)) \geq s \cdot (1 + \delta) a_j \\ U_B(a_j) &= v_L + a_j r \geq a_j s \end{aligned}$$

Where, because good politicians are reelected their utility of being in office has to outweigh their wage for two periods, while since bad politicians are not re-elected, their utility only has to be higher than their wage for a single period. Solving these equations for a_i pins down the marginal types:

$$\begin{aligned} a_{G,1}^* &= \frac{v_L + \delta(\rho v_N + (1 - \rho) v_L)}{(1 + \delta)s + \bar{x}(\gamma) (g - u) - g (1 + \delta(1 - \rho))} \\ a_{B,1}^* &= \frac{v_L}{s - r} \end{aligned}$$

Proposition 5. *Higher quality local news improves selection into politics in period 1, both in terms of average ability, as well as the probability that a politician is the good type. Selection*

into the pool of challengers in period 2 is unaffected

Proof. Period 2 pay-offs are not affected by the quality of local news, since there is no more accountability and politicians implement the policies that maximize their utility independent of what beliefs this might induce in either voters or parties.

The marginal Good type who decides to run in period 1 is increasing in γ , since $\bar{x}'(\gamma) < 0$. Bad types are always corrupt and are never reelected or promoted, so their entry decision is unaffected by the quality of local news. An increase in the quality of local news thus affects the marginal type willing to run for office ($(a_{G,1}^*)'(\gamma) > 0$), the average quality of the good politician ($E[a|a \leq a_{G,1}^*]$) as well as the mass of good politicians running for office ($F(a_{G,1}^*)$). \square

Both parties and voters benefit from improved selection. National parties because the politicians they promote are of higher ability. Local voters because they get better policy pay-offs. At the same time, there is once again a tension: where career incentives make voters better off in period 1, in period 2 this is only the case if the politician sticks around. While it might be optimal from a local perspective for the politician to be promoted (that will be discussed in further detail below), from a local perspective a promotion does not just mean throwing away information about the type, but also that the average politician that will replace them will be of lower quality.

In the next section, I will investigate whether, given these costs and benefits, local voters are willing to elect career politicians at all, or whether they would rather elect outsider candidates or politicians from local parties, who do not face career incentives at all.

5.2 Local versus national parties

Given that career incentives lead to a distortion in behavior of good politician, an obvious follow-up question to ask is whether voters would actually be interested in electing a politician who faces career incentives. The low approval rating of national political parties, the rise of local parties and independent candidates suggests that the answer is not an unequivocal yes: at least in some cases, the electorate seems to prefer politicians who have no chance of ever running for higher office. In the next section, I argue that one contributing factor for this trend is the decline in local news because it increases the distortion caused by career incentives. I do this by extending the baseline model by adding a stage to the game where voters get to choose between electing a politician with career incentives and one without.

What exactly is the role of national political parties in local politics? Here I will build on the insights from the previous section. I will not use the full selection model explored in Section 5.1 and Appendix B, but will rely on a simplified reduced form, in which national parties offer career incentives and attract higher ability politicians than would usually run for local office in period 1. Voting for a national party has two benefits: the initial pool of candidates is of higher ability, and, because good politicians follow the party line, they provide information about the type of the politician before the first election. It comes at the cost of party line voting in period 1 by good politicians as well as losing the politicians to a promotion before period 2.

What about local parties? These will be modeled as parties only active at the local level, and can also be thought of as independent candidates or local figures like business owners who run

for office independently from existing parties or endorsements from national politicians. What do they offer voters? Two things. First, I will assume that they do not pick the uniform policy to signal their type. So voters can be guaranteed to get bottom-up policy, although not whether the politician will be corrupt or not. This is consistent with promises of these types of politicians to not listen to ‘elites in the national capital.’ This also means the behavior will be very similar to that seen in standard games in the literature. Second, because they are not part of a national party, they commit to staying in local politics. This means that *if* the politician turns out to be the good type, they can always be reelected.

Players

There are two types of parties: National parties and Local parties, indicated by $\{L, N\}$. National parties care about finding good candidates for national office. Local parties care about winning office.

$$P_N = \begin{cases} a & \text{If they promote a good politician of ability } a. \\ 0 & \text{If they promote a bad politician.} \end{cases}$$

$$P_L = \begin{cases} 1 & \text{If they win local office.} \\ 0 & \text{If they do not.} \end{cases}$$

Because they have a chance of using local office as a stepping stone for national office, good politicians who select into office for the National party in the first period are of ability $a(1 + \gamma)$, with $a > 1$. All other politicians, including bad politicians, politicians who run in the second period, those who only seek a promotion and those who select into the Local party, are of ability normalized to 1. I assume that the probability of selecting a good politician is constant, so $P(T = G) = \pi$, independent of party and ability.

Timing

Period 0

1. Politicians of ability $a(1 + \gamma)$ select into running for the National party. Politicians of ability $a = 1$ select into running for the Local party.
2. There is an election in which voters choose between electing either the Local or the National party, choosing whichever party maximizes their expected utility.

Period 1 & 2

- Identical to the earlier game if a national party is selected.
- If voters choose a local party, there is no promotion opportunity and only an election.
- All challengers are of ability $a = 1$.

Solution

National party

The behavior of the national party is equal to what was derived in Section 4 above. Because there is positive selection of talent into local politics in period 1 (since $a(1 + \gamma) > 1$ is larger than 1 if $a > 1$), conditional on learning that the type of the politician is good, parties want to promote these politicians and voters want to reelect them.

The welfare of choosing a national party is then equal to:

$$V(\text{Party} = N, \gamma) = \pi \left[a(1 + \gamma) \cdot (g - \bar{x}(\gamma) (g - u)) + \delta \cdot g \cdot (a(1 + \gamma) (1 - \rho) + \pi \rho + (1 - \pi)) \right] \quad (2)$$

Note that voters only benefit from electing a high ability politician if they are of the good type (and thus implement good policies, rather than simply being corrupt) and as long as they are not promoted. Once a politician is promoted, local voters have to elect a news politician, who will be of ability 1.

A decline in local news has two effects. First, it increases the amount of top-down policy making in period 1, and second, it decreases the average ability of politicians who run for office in period 1.

Local Party

In period 2, behavior is the same as in the equilibrium describe above. Good politicians choose the good policy for all public goods. Bad politicians choose the bad policy.

To make the strongest and most interesting comparison between the local and the national party, I impose that politicians from the local party do not use u to signal their type. This rules out that a politician from the local party behaves like a slightly worse ability version of a politician from the national party.

Assumption 4. *Politicians from the local party do not choose u in equilibrium.*

Then if voters and parties do not expect any uniform policy making from politicians, but rather expect good politicians to only choose good policies and bad politicians only choose bad policies. The newspaper is then informative about the choice of the politician. Beliefs of both parties and voters are then:

$$\hat{\pi}_P = \hat{\pi}_V = \begin{cases} 1 & \text{If the newspaper reports } g \text{ for all policies.} \\ \pi & \text{If there no newspaper report.} \\ 0 & \text{Otherwise.} \end{cases}$$

Good politicians have no reason to deviate and choose $\{x^u = 0, x^g = 1, x^b = 0\}$. Choosing anything else reduces their first period pay-offs, while not increasing their probability of promotion or reelection. In this equilibrium, bad politicians choose $\{x^u = 0, x^g = 0, x^b = 1\}$ and are reelected and promoted if there is no news report, and lose their position if there is. They

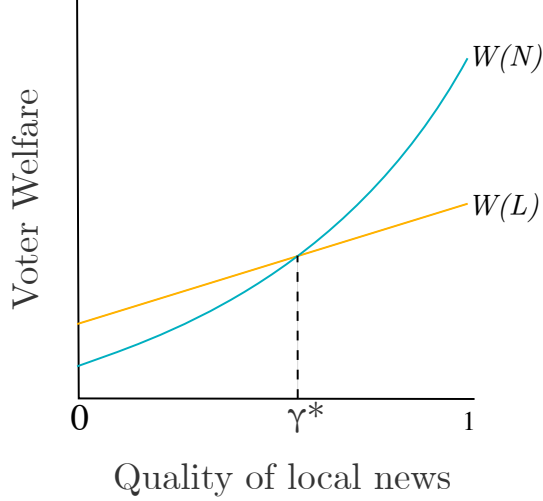


Figure 6: Utility from both local and national parties is declining in the quality of local news.

could try to increase their pay-offs by deviating and mimicking the good politician's behavior. If Assumption 1 holds, so if rent is sufficiently valuable, they never want to do this.⁹

Utility of voting for the local party is then:

$$V(\text{Party} = L, \gamma) = \pi g\left(1 + \delta(1 + (1 - \pi)\gamma)\right) \quad (3)$$

A decline in local news has a single effect: it reduces period 2 selection, as bad politicians who pursue corrupt projects are less likely to be discovered and more likely to get re-elected.

Voters

The question is then whether voters get more welfare if they vote for the local party or the national party. This comes down to a comparison between Equations 2 and 3. The trade-off is as follows. The national party attracts higher talent, and through party line voting, information is revealed about the type of the politician. This comes at the cost of policy adaptation in period 1 and the chance of losing good politicians to a promotion in period 2. The local party on the other hand attracts worse quality politicians in period 1, but does not constrain their behavior and voters can be sure that if they want to keep the politician, they can, since there is no chance of promotion.

Voters prefer voting for a national party if:

$$V(N, \gamma) \geq V(L, \gamma) \quad (4)$$

First, note that the utility from voting for a national party depends on $\bar{x}(\gamma)$, which is a linear

⁹To see this, note that since $v_N > v_L$, $r \geq (\delta(\rho v_N + (1 - \rho)v_L))/(1 - \delta(1 - \rho))$ implies $r > (\delta v_L)/(1 - \delta)$. This means this (partial information) equilibrium for the local party coexists with the topdown equilibrium for the national party.

function of γ . To save on notation, it is useful to define:

$$\begin{aligned}\bar{x}(\gamma) &= (1 - \gamma) \frac{\delta}{r} [\rho v_N + (1 - \rho)(r + v_L)] \\ &\equiv (1 - \gamma) D(\delta, r)\end{aligned}$$

Where the $D(\delta, r)$ term captures the top-down distortion in the behavior of politicians from the national party. This allows me to rewrite Equation 4 as the following quadratic equation:

$$\begin{aligned}\gamma^2 \left(a \cdot (g - u) D(\delta, r) \right) + \gamma \left(g(a + \delta((1 - \rho) a - (1 - \pi))) \right) + \\ \left(g(a - 1) - a \cdot (g - u) D(\delta, r) + \delta g(a + \pi)(1 - \rho) \right) \geq 0\end{aligned}\tag{5}$$

There are three main drivers that determine the sign of this equation.

Where the first term captures the dual benefit of improving the quality of local news when voting for the national party: it improves local selection of ability, and improves first period behavior, by giving good politicians of national parties more room to pursue good rather than uniform policies. The second term captures that while better local news attracts higher local candidates for the national party, the local population only benefits if in period 2 if there is no promotion (which happens with probability $1 - \rho$). At the same time, an increase in the quality of local news improves second period selection for the local party, so if they picked a bad candidate (which happens with probability $1 - \pi$) the newspaper helps reveal this. The final term captures the static benefits, which captures the trade-off the attraction of higher ability politicians with the benefits of policy adaptation.

Lemma 4. *National parties are preferred for high quality local news if they attract sufficiently high ability politicians.*

Proof. If $\gamma = 1$ there is no distortion in the behavior of the national politicians, while bad politicians from the local party are always discovered and thus never reelected. The higher ability that the national party attracts then only has to outweigh the cost of promotion:

$$a \geq \frac{1}{2} \frac{1 + \delta(1 - \pi\rho)}{1 + \delta(1 - \rho)}$$

Note that if $\rho = 0$, this always holds. □

If the quality of local news is high, voting for the national party has no cost in terms of party line voting, but it does come at the cost of potentially losing a high ability politician if there is a promotion opportunity. The ability of the politician the party attracts has to be sufficiently high to compensate for this cost. If there are no promotions (so if $\rho = 0$) then this only imposes that $a(1 + \gamma)|_{\gamma=1} > 1$: positive selection into the national party.

Lemma 5. *Local parties are preferred for low quality local news if the value of the good policy is sufficiently high.*

Proof. Evaluating Equation 5 at $\gamma = 0$:

$$g(a - 1) - a \cdot (g - u)D(\delta, r) + \delta g(a + \pi)(1 - \rho) \geq 0$$

$$g \geq \frac{a \cdot D(\delta, r) \cdot u}{a \cdot D(\delta, r) - (a - 1) - \delta(a + \pi)(1 - \rho)}$$

□

This condition captures the key reason why voters would ever elect a Local party. When voting for a National party, good politicians will follow the party line in period 1 in order to induce positive beliefs. For those policies, voters get the uniform policy rather than the good policy. Voters prefer voting for a local party if period 1 policy adaptation is sufficiently valuable to them. This will of course depend on the policy pay-offs u and g . But it also depends on how much good politicians have to tow the party line: the larger this distortion (captured by $D(\delta, r)$), the more likely voters prefer voting for a Local party. Finally, the worse the ability of the politicians the national party attracts, the more likely voters prefer a local party.

Proposition 6. *Suppose the conditions derived in Lemmas 4 and 5 hold. Then if the quality of local news is high ($\gamma > \gamma^*$), voters prefer to elect the national party. If the quality of local news is below this threshold, they prefer a local party.*

Proof. Given Lemmas 4 and 5, the National Party is preferred if $\gamma = 1$ and the Local party is preferred if $\gamma = 0$. Since the indifference equation is quadratic, convex in γ and increasing between $\gamma \in [0, 1]$, the largest of the two solutions will be internal.

Given that this is a quadratic equation, it has the standard solution:

$$\gamma^* = \frac{\sqrt{(f_2)^2 - 4 \cdot f_1 \cdot f_3} - f_2}{2 \cdot f_1}$$

Where:

$$f_1 = a \cdot (g - u)D(\delta, r)$$

$$f_2 = g(a + \delta((1 - \rho) a - (1 - \pi)))$$

$$f_3 = g(a - 1) - a \cdot (g - u)D(\delta, r) + \delta g(a + \pi)(1 - \rho)$$

□

National parties attract higher ability candidates, but in equilibrium, good politicians of these parties choose to follow the party line. This is costly for voters. First, because they get lower policy pay-offs in period 1 (u rather than g) but also because this distortion is coming from the good types, who would in the absence of career incentives implement be naturally inclined to implement good policies. Second, because the information revealed is only beneficial if the voters get to keep the politician. This is why, in this model, local parties are attractive: in the case voters happen to elect a good politician, they will be unhampered in implementing what is locally most beneficial to local voters. In fact, local parties are preferred for low quality local news, despite their politicians being not just of worse average ability, they are also more likely to be corrupt:

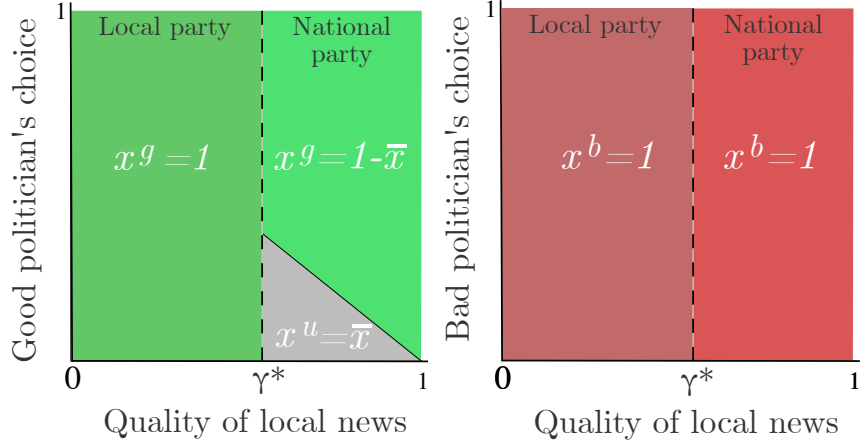


Figure 7: The behavior of good and bad politicians in period 1 as a function of the voter's choice for the local or the national party. Bad politicians always behave the same. Good politician's of the local party distort their behavior to signal their type.

Corollary 5.1. *If the quality of local news is low ($\gamma < 1 - \rho$), politicians of local parties are more likely to be the bad type than those of national parties.*

Proof. In period 1, politicians of both parties are equally likely to be the bad type ($1 - \pi$).

In period 2, politicians from the local party are the bad type if a bad politician is elected in period 1 and not discovered, or if they are discovered and the challenger is also bad. The combined probability is $(1 - \pi)(1 + \gamma(1 - \pi))$.

For the national party, bad types reveal their type through the action. A bad politician can be elected in period 2 if (1) a good politician is selected in period 1 and promoted and replaced by a bad challenger (2) a bad politician is selected in period 1 and replaced by a bad challenger. The combined probability is: $\pi(1 - \pi)\rho + (1 - \pi)^2$.

$$\begin{aligned}
 P(T = B|t = 2 \& P = L) &\geq P(T = B|t = 2 \& P = N) \\
 \pi(1 - \pi)\rho + (1 - \pi)^2 &\geq 1 + \gamma(1 - \pi) \\
 \gamma &\leq 1 - \rho
 \end{aligned}$$

If $\gamma = 0$ this will be true since $\rho < 1$. □

This makes clear how costly party line voting can be. Local voters might prefer to take the risk of reelecting a bad politician over good politicians distorting their behavior in period 1.

Corollary 5.2. *If the conditions for Proposition 6 are met, there is a non-monotonicity in the amount of bottom-up policy making by good politicians.*

- *If $\gamma = 1$: The national party is elected and good politicians do not to choose u in equilibrium ($\bar{x}(\gamma = 1) = 0$).*
- *If $\gamma \in (\gamma^*, 1]$: The national party is elected, but good politicians have to follow the party line in order to signal their type. The lower the quality of local news, the more uniform policy making is expected of them.*

- If $\gamma \in [1, \gamma^*)$: Voters elect the local party and politicians no longer choose u .

This proposition is the result of the differences between local and national parties. When the quality of local news is very high, national parties can give a lot of freedom to their local candidates and still be confident that they are only promoting good types. As the quality of local news declines, they require more top-down policy making from their local representatives. The cost of this is borne by politicians and voters, not by the parties. When this cost becomes too high, voters start electing candidates who use full discretion, even if they learn less about the politician and are more likely to re-elect bad politicians this way.

5.2.1 Discussion

Career incentives are often highlighted as a way to improve the behavior of local politicians. But career incentives also come with costs, which, at least in this model, are borne by local voters rather than by parties. These costs might be so high that voters prefer electing local parties, who cannot offer career incentives. Even if local parties attract lower ability politicians, and are (over the course of the two periods) more likely to be corrupt. This nuances the arguments made in Myerson (2006). He argues that local politicians have a reason to behave well in order to try to secure a job in national politics. What I argue here is that voters might try to find a way to elect politicians that reject this path, severing the pipeline between national and local politics. In this way, it implies that the types of parties that arise at the local level, and the incentives they offer, is endogenous to the quality of local news.

5.3 Federalism and decentralization (work in progress)

By definition, local government is not the only game in town. Policy can always be made in the capital. This section explores, given the dynamics above, how a decline in local news affects the rationale for decentralization.

The usual trade-off considered in models of fiscal federalism is one between saving tailoring policy to local needs and saving on fixed costs. Decentralizing has the benefits that local politicians can use their local information to adapt policy to local circumstance. But it comes at a cost: every region needs to invest in the local bureaucracy and capacity to actually implement these policies. Centralization has the benefit of only having to make that investment once, thus saving on fixed costs and capitalizing on economies of scale.

This model adds three additional forces. First, local politicians might not have the incentive to use local information. This reduces the benefits of decentralization, because not all the benefits of decentralization are reaped. Second, party line voting produces information and improves national selection. Finally, local voters might vote for a local party and shut off the pipeline between local and national politics.

I will compare two scenarios, one if national parties are elected, and one if local parties are elected. Given that all regions are assumed to be identical, all regions have the same quality local newspaper and elect the same type of party. In reality, regions will differ along many dimensions.

5.3.1 Set-up

There is a continuum of regions. Each of these is assumed to be identical on observables, but differ in (nationally) unobservable ways that create opportunities for local policy adaptation.

There is a continuum of public good $x_i \in [0, 1]$. These are provided by either the local or the national government. A fraction $x < 1$ of these policies are implemented by the local government, the remainder $1 - x$ is implemented by the national government.

The quality of policy making depends on the stock of talent at the local and the national level in office that period. The local stock of talent will be modeled as in the previous section. Define $a_{N,t}$ as the stock of talent available at the national level, which depend on the ability and types of politicians in office that period. $a_{N,1}$ is exogenous. After period 1, a share of $\rho\pi$ politicians retire, and $a_{N,2}$ is determined by their replacements.

Although centralization allows for policy adaptation, it is also costly: instead of pooling resources, it requires that each region has its own bureaucracy. I assume that the cost of decentralization is: $c(x) = \frac{1}{2}x^2$. The optimal level of decentralization, x^* , will be chosen by a benevolent social planner before the game starts. The social planner's problem is then:

$$W(x) = \int_0^1 \left(a_{j,1}(\gamma) \int_0^x w(x_{i,1}) di + a_{j,2}(\gamma) \delta \cdot \int_0^x w(x_{i,2}) di \right) dj + (1 - x) u(a_{1,N} + \delta \cdot a_{N,2}(\gamma)) - c(x)$$

Timing

Period 0

1. The social planner chooses the optimal level of decentralization x^* .
2. Voters choose whether to elect a local or a national party.

Period 1 and 2

Identical as before, but a share $\rho \cdot \pi$ national seats open up. If a national party is elected at the local level, politicians who are promoted occupy those seats. If a local party is elected, a random draw from the pool of politicians is taken.

5.3.2 Solution

I will compare two scenarios, one if national parties are elected, and one if local parties are elected.

If a national party is elected

Local politician's behavior

Any generic local politician takes the level of decentralization x as given. Period 2 behavior is similar to what was derived before, good politicians make good policy for all policies under their control $\{x^u = 0, x^g = x, x^b = 0\}$, bad politicians make bad policy $\{x^u = 0, x^g = 0, x^b = x\}$. Note that this is exactly what they did in the main model, replacing 1 with x .

In period 1, level of party line voting now depends on x as well. This is for straightforward reasons: the more decentralized policy making is, the more attractive it is to be corrupt for bad politicians. In equilibrium, top-down voting has to adjust to account for this.

$$\begin{aligned} U_B(x^u = 0, x^g = 0, x^b = x) &\geq U_B(x^u = \bar{x} \cdot x, x^g = 0, x^b = (1 - \bar{x}) \cdot x) \\ rx + v_L &\geq (1 - \bar{x}) rx + v_L + \delta(1 - \gamma)(\rho v_N + (1 - \rho)(rx + v_L)) \end{aligned}$$

Solving this gives:

$$\bar{x}(\gamma, x) = \frac{1}{rx} (1 - \gamma)(\rho v_N + (1 - \rho)(v_L + rx)) \quad (6)$$

Since there is a continuum of regions, the stock of local talent is simply equal to the average level of talent times the probability of selecting a good type.

$$\begin{aligned} a_{L,1}(\gamma) &= \pi a(1 + \gamma) \\ a_{L,2}(\gamma) &= \pi (a(1 + \gamma)(1 - \rho) + \pi \rho + (1 - \pi)) \end{aligned}$$

National politician's behavior

National politician's can only choose u . The marginal benefits of national policy making depend on the stock of national talent available in that period. In period 1 the stock of talent $a_{N,1}$ is given. Before period 2, a share $\rho\pi$ seats open up. This is consistent with every good politician believing they have a probability ρ of being promoted. The stock of talent in period 2 is then equal to:

$$\begin{aligned} a_{N,2}(\gamma) &= (1 - \rho\pi) a_{N,1} + \rho\pi a_{N,1} \\ &= (1 - \rho\pi) a_{N,1} + \rho\pi a(1 + \gamma) \end{aligned}$$

The higher the quality of local news, the better selection into local politics in period 1, improving the local talent pool. This improves national selection in period 2 as these politicians get promoted.

The optimal level of decentralization

The problem of the social planner is to maximize total social welfare, taking into account how politicians will actually behave.

$$W(x) = x \cdot \left(a_{L,1}(\gamma)(g - \bar{x}(x, \gamma)(g - u) + \delta \cdot a_{L,1} \cdot g) \right) + (1 - x)(a_{1,N} + \delta \cdot a_{2,N}(\gamma)) \cdot u - c(x)$$

What if there are no distortions in behavior of the good type, so if $\bar{x} = 0$? That gives a first best benchmark.

$$x_{FB}^* = \frac{g \cdot (a_{L,1}(\gamma) + a_{2,N}(\gamma) \cdot \delta) - u \cdot (a_{N,1} + a_{N,2}(\gamma) \cdot \delta)}{c}$$

Then the optimal level of decentralization depends on three factors. First, the relative benefits

of policy adaptation ($g-u$). Second, the stock of talent of local and national politicians. Finally, the cost parameter c .

If there are distortions in behavior of local politicians. Which has the following first order conditions:

$$a_{1,L}(\gamma) \cdot \left(g + \left(\bar{x}(x^*) - \frac{d\bar{x}(x^*, \gamma)}{dx} \right) (g-u) \right) + \delta \cdot a_{2,L}(\gamma) \cdot g - (a_{1,N} + \delta \cdot a_{2,N}(\gamma)) \cdot u = c \cdot x^*$$

Relative to the first order conditions of first best, there is one additional term. This is because the amount of bottom-up policy making in period 1 is now no longer equal to x^* , but rather $x^* \cdot (1 - \bar{x}(x^*, \gamma))$.

Proposition 7. *If a politician from the national party is elected, it is the political party and the career incentives they offer that determine the de facto level of decentralization in period 1. The de jure level of decentralization x only has an indirect effect, by changing these incentives.*

Proof. Let x be the de jure level of decentralization. Define the de facto level of decentralization as the amount of policies for which good local politicians actually use local information.

In period 2, good politicians can use local information for all policies, since they have no more reason to signal their type.

In period 1, good politicians only use local information for only a share of policies $x^* (1 - \bar{x}(\gamma))$. \square

What this proposition captures is an old idea within the federalism literature, which goes back at least to Riker (1964): even if constitutionally local politicians have a lot of power, it is the political system and the political parties that determine how decentralized a political system actually is. Holding the de jure level of decentralization x^* constant, a decline in the quality of local news leads to a *de facto* centralization, with local politicians choosing not to adapt policies and instead doing what national elites would have done.

This is also reflected in the optimal level of decentralization given distortions in behavior:

$$x^* = \frac{g \cdot (a_{L,1}(\gamma) + a_{L,2}(\gamma) \cdot \delta) - u \cdot (a_{L,1}(\gamma)(1-\gamma)(\rho v_N + (1-\rho)v_L) + (a_{N,1} + a_{N,2}(\gamma) \cdot \delta))}{c + 2 \delta \cdot a_{L,1}(\gamma) \cdot (1-\gamma)(g-u)}$$

The optimal level of decentralization is now affected by the amount of party line voting of local politicians. This affects both the benefits of decentralization. The benefits of decentralization are lower, because local politicians implement the same policies that national politicians would.

A decline in the quality of local news affects the optimal level of decentralization through multiple channels. First, it reduces the amount of policy adaptation that good politicians engage in. This reduces period 1 policy pay-offs at the local level. Second, it affects the average ability of politicians. This is true both at the local level (in both periods), but it also reduces the average quality of politicians promoted to the national level in period 2. If local politics fails to attract good politicians, and local politics serves as a springboard for a national career, a reduction in local talent also makes the national government less effective.

If a local party is elected

Electing local parties has three effects. First, behavior at the local level will be different. Politicians no longer follow the party line. Second, there is less information generated. Third, the pipeline between local and national politics is shut down.

Local politicians

Local behavior is very similar to what was derived before. In both periods, good politicians choose good policies for all policies under their control ($\{x^u = 0, x^g = x, x^b = 0\}$) and bad politicians choose bad policies for all policies under their control ($\{x^u = 0, x^g = 0, x^b = x\}$). Bad politicians are discovered and replaced with probability γ .

National politicians

National politicians choose the uniform policy u . In period 1, the stock of national talent is given as $a_{N,1}$. In period 2, all politicians who retire are replaced by a random draw from the pool of potential politicians.

$$a_{N,2} = a_{N,1} - \rho\pi (a_{N,1} - \pi)$$

The optimal level of decentralization

If local parties are elected in the regions, then the social planner's objective is equal to:

$$W(x) = x \cdot (g \cdot \pi (1 + \delta(1 - \pi)\gamma)) + (1 - x) \cdot u \cdot (a_{N,1} + \delta \cdot a_{N,2}) - c(x)$$

Which was the following solution.

$$x^* = \frac{g \cdot \pi (1 + \delta(1 - \pi)\gamma) - u \cdot (a_{N,1} + \delta \cdot a_{N,2})}{c}$$

In this case, the optimal level of decentralization is determined by just the relative effects of local and national policy making. There are no spillovers between the layers of government.

The quality of local news reduces the optimal level of decentralization. It does this by affecting second-period selection at the local level: lower quality local news means that politicians are more likely to get re-elected despite being corrupt. This means that a decrease in the quality of local news also decreases the optimal level of decentralization.

Beyond that, there is one more important difference with the case where a national party is elected. By breaking the link between local and national politics, local politicians no longer make the jump to national politics. This reduces the stock of talent available to make policy in the second period.

6 Conclusion

In this paper I explored the national implications of the decline of local news. First, I investigate how career and electoral incentives are changed, focussing on how it distorts the behavior of

good types: those motivated by making optimal policy decisions. I show that good types can effectively signal their type by not adapting policy, but rather sticking to the national party line. The range of policies for which politicians have to follow the party line and still effectively signal their type increases as the quality of local news declines. This reduces welfare for voters not through abuse of office, but rather an increase in the underuse of the power of holding local office.

I then develop three extensions. First, I argue that the more good politicians have to distort their behavior, the less attractive it is for them to hold local office. This reduces selection at the local level. Second, I give voters the choice between voting for a national party, which attracts higher ability candidates, but where good politicians follow the party line, or a local party, of worse ability candidates who always choose to make full use of local office. If the quality of local news is high, voters elect national parties which give complete freedom to their local politicians. If the quality of local news is very low however, the cost of party line voting becomes too high and voters switch to voting for local parties. This is true, even if politicians from local parties are more likely to be corrupt. Finally, I place these findings in a simple federalism framework and endogenize the optimal level of decentralization. High quality local news improves both local and national politics. Local politics because it means that politicians make full and optimal use of local office. National politics because improved selection of politicians in period 1 at the local level creates a higher quality talent pool for national politics in period 2.

References

- Alesina, A. and E. Spolaore (2003). *The Size of Nations*. Mit Press. Cambridge and London: MIT Press.
- Angelucci, C. and J. Cagé (2019). Newspapers in times of low advertising revenues. *American Economic Journal: Microeconomics* 11(3), 319–364.
- Arons, T. M., J. Hoppock, and M. Arons (2008, feb). Obama as a State Senator: Does He ‘Deserve a Promotion?’.
- Barber, B. R. (2013). *If Mayors Ruled the World: Dysfunctional Nations, Rising Cities*. Yale University Press.
- Besley, T. (2006). *Principled agents?: The political economy of good government*. Oxford University Press, USA.
- Borchert, J. (2009). Ambition and opportunity in federal systems: the political sociology of political career patterns in Brazil, Germany, and the United States. In *APSA 2009 Toronto Meeting Paper*.
- Cagé, J. (2020). Media competition, information provision and political participation: Evidence from French local newspapers and elections, 1944–2014. *Journal of Public Economics* 185, 104077.

- Cho, I.-K. and D. M. Kreps (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics* 102(2), 179–221.
- Dal Bó, E., F. Finan, O. Folke, T. Persson, and J. Rickne (2017). Who Becomes A Politician? *The Quarterly Journal of Economics* 132(4), 1877–1914.
- Doctor, K. (2015). The newsonomics of how and why.
- Drago, F., T. Nannicini, and F. Sobbrío (2014). Meet the press: How voters and politicians respond to newspaper entry and exit. *American Economic Journal: Applied Economics* 6(3), 159–188.
- Enikolopov, R. and E. Zhuravskaya (2007). Decentralization and political institutions. *Journal of public economics* 91(11-12), 2261–2290.
- Fowler, L. L. and R. D. McClure (1989). *Political Ambition: Who Decides to Run for Congress*. Yale University Press.
- Gao, P., C. Lee, and D. Murphy (2020). Financing dies in darkness? The impact of newspaper closures on public finance. *Journal of Financial Economics* 135(2), 445–467.
- Gavazza, A., M. Nardotto, and T. Valletti (2019). Internet and politics: Evidence from UK local elections and local government policies. *The Review of Economic Studies* 86(5), 2092–2135.
- Hamilton, J. T. (2016). *Democracy’s Detectives: The Economics of Investigative Journalism*. Harvard University Press.
- Hopkins, D. J. (2018). *The increasingly United States: How and why American political behavior nationalized*. University of Chicago Press.
- Keller, B. (2015, apr). David Simon on Baltimore’s Anguish.
- Long, R. (2007, may). State pork to Obama’s district included allies, donors.
- Lorentzen, P. (2014). China’s strategic censorship. *American Journal of political science* 58(2), 402–414.
- Mair, P. (2013, aug). *Ruling The Void: The Hollowing Of Western Democracy*. London ; New York: Verso.
- Martinez-Vázquez, J., S. Lago-Peñas, and A. Sacchi (2017). The impact of fiscal decentralization: A survey. *Journal of Economic Surveys* 31(4), 1095–1129.
- Moskowitz, D. J. (2021). Local news, information, and the nationalization of US Elections. *American Political Science Review* 115(1), 114–129.
- Myerson, R. B. (2006). Federalism and Incentives for Success of Democracy. *Quarterly Journal of Political Science* 1, 3–23.
- Nielsen, R. K. (2015, aug). *Local Journalism: The Decline of Newspapers and the Rise of Digital Media*. I.B.Tauris.

- Oates, W. E. (1972). *Fiscal Federalism*. Harbrace series in business and economics. Harcourt Brace Jovanovich.
- Obama, B. (2020). *A Promised Land*. Penguin Books Limited.
- Richards, S. (2017). *The rise of the outsiders: How mainstream politics lost its way*. Atlantic Books.
- Riker, W. H. (1964). *Federalism : Origin, Operation, Significance*. Basic studies in politics. Little, Brown.
- Riker, W. H. (1975). Federalism. In F. Greenstein and N. Polsby (Eds.), *Governmental Institutions and Processes*, Chapter 2, pp. 93–172. Reading, MA: Addison-Wesley.
- Simon, D. (2009). Testimony, U.S. Senate Commerce Committee, Hearing on The Future of Newspapers.
- Tabellini, G. (2005). The role of the state in economic development. *Kyklos* 58(2), 283–303.
- Trussler, M. (2020). Get information or get in formation: The effects of high-information environments on legislative elections. *British Journal of Political Science*, 1–21.
- Walker, M. (2021, jul). U.S. newsroom employment has fallen 26% since 2008.

A All equilibria

This game has three types of equilibria: Full information separating equilibria (of which the one described in the main text is the most plausible), pooling equilibria (which do not exist if the full information separating equilibrium exists) and partial information separating equilibria.

Fully informative separating equilibria

Note that the conditions to prove Proposition 2 are inequality constraints rather than equality constraints. In other words, there are other equilibria with identical belief structure, but rather than x^u being equal to the minimal level of top down policy making derived in Prop 1, the minimal level is some $\bar{x}' \in (\bar{x}, 1]$. Bad politicians then always choose to be corrupt, while good politicians choose $\{x^u = \bar{x}', x^g = 1 - \bar{x}', x^b = 0\}$. This is an equilibrium under the following beliefs:

$$\hat{\pi}_P = \hat{\pi}_V = \begin{cases} 1 & \text{If } x^u \geq \bar{x}' \text{ and the newspaper did not find abuse of office.} \\ 0 & \text{Otherwise.} \end{cases}$$

Note that although this does not impose further restrictions on the bad type (they would surely not deviate), it does increase the costs for good types.

Pooling

In a pooling equilibrium, both types of politicians make the exact same choice in period 1. Their choice is thus uninformative, and the newspaper has nothing to report. Both types of politicians are re-elected and promoted, although voters and parties are less informed when they make their choice.

There is an infinite number of these equilibria, which depend on the beliefs of voters and parties. For a fixed $\bar{x}' \in [0, \bar{x})$, politicians choose policies $\{x^u = \bar{x}', x^g = 1 - \bar{x}', x^b = 0\}$ in period 1. As long as the newspaper does not report any abuse of office and the politician chooses at least $x^u \geq \bar{x}'$, voters believe the politician is the average type.

$$\hat{\pi}_P = \hat{\pi}_V = \begin{cases} \pi & \text{If } x^u \geq \bar{x}' \text{ and the newspaper did not find abuse of office.} \\ 0 & \text{Otherwise.} \end{cases}$$

There are two things to note. First, note these equilibria require that bad politicians do not choose bad policies in period 1, but rather choose to induce positive beliefs and get re-elected or promoted in period 2. This means that this equilibrium does not coexist with the equilibrium described above. In other words, this equilibrium exists only if corruption rents are relatively low, so if Assumption 1 does not hold.

Second, note that almost all of these equilibria are implausible. Politicians only choose the uniform policy because it is expected of them, it does not benefit voters (who would prefer g for these policies) nor parties (who learn nothing from this choice). The most plausible equilibrium is the one in which $\bar{x}' = 0$ and politicians choose the good policy for all policies in the first period. Then behavior is exactly the same as one would see in the pooling equilibrium described in Besley (2006).

Partially informative separating equilibria

Finally, there are separating equilibria in which the level of expected top-down policy making is not sufficiently high to deter bad politicians from mimicking good politicians. At the same time, corruption is sufficiently attractive that bad politicians simply partially mimic good politicians and hope the newspaper does not discover them. This includes no top-down policy making expected at all. This equilibrium is then equivalent to the standard separating equilibrium found in these types of games.

In these equilibria, good politicians choose $\{x^u = \bar{x}', x^g = 1 - \bar{x}', x^b = 0\}$, with $\bar{x}' \in [0, \bar{x})$. Bad politicians choose $\{x^u = \bar{x}', x^g = 0, x^b = 1 - \bar{x}'\}$. The newspaper is informative: if it reports wrongdoing, voters and parties know the politician's type is bad. If the newspaper reports that the politician chose the good policy, they know the politician is the good type. Since no report is not informative, if there is no report voters and parties stick with their prior belief.

$$\hat{\pi}_P = \hat{\pi}_V = \begin{cases} 1 & \text{If } x^u \geq \bar{x}' \text{ and the newspaper reports } g \text{ for the remaining policies.} \\ \pi & \text{If } x^u \geq \bar{x}' \text{ and there is no newspaper report.} \\ 0 & \text{Otherwise.} \end{cases}$$

Note that, like in the previous case, most of these equilibria are sustained by implausible beliefs. The uniform policy is only chosen because it is expected of politicians, not because it reveals any information. The most plausible alternative is $\bar{x}' = 0$, then good politicians choose good policies, bad politicians bad policies and newspaper coverage determines whether bad politicians get away with it.

In this equilibrium, the pay-offs for politicians are highest if $\bar{x}' = 0$ (because then neither type has to give up any policy or corruption rents, while still (having a shot at) being promoted and re-elected). Pay-offs are then equal to:

$$U_G = g + v_L + \delta(\rho v_N + (1 - \rho)(v_L + g))$$

$$U_B(\gamma) = r + v_L + (1 - \gamma)(\rho v_N + (1 - \rho)(v_L + g))$$

Discussion

Although there is an infinite number of equilibria, most of these are implausible or do not exist when the equilibrium derived in the main text exists. Fully informative separating equilibria require ‘too much’ party line voting. Pooling equilibria do not exist if the topdown equilibrium exist. The most plausible alternative is the partially informative separating equilibrium, where no party line voting is expected of politicians. I will now derive the condition found in the main text under which the topdown down equilibrium is the only equilibrium that remains after refinement using the Intuitive Criterion.

A.1 Equilibrium refinement

Without any further assumption, both the topdown equilibria and the partial information equilibria coexist if Assumption 1 and 2 hold.

A.1.1 Full information equilibria

For convenience, let me restate that in Assumption 3 I impose that the value of a promotion $v_N(\hat{\pi}_P)$ is now a function of the beliefs of the party.

I will now prove that the topdown equilibrium presented in the text is the only one that survives refinement by the Intuitive Criterion.

First, note that the minimum level of x^u now depends on the beliefs of the party:

$$\bar{x}(\gamma) \geq (1 - \gamma) \frac{\delta}{r} (\rho v_N(\hat{\pi}_P = 1) + (1 - \rho)(r + v_L))$$

Given this ‘new’ level of minimum expected party line voting, can good types convince parties that only they would deviate to the full information equilibrium?

Note that although there are an infinite number of fully informative equilibria, in most of these good politicians go above and beyond what they need to do to prove their type (so where they are expected to choose u for $x^u > \bar{x}$). It is thus also clear that these equilibria do not meet the Intuitive Criterion. Good types would benefit from deviating to $x^u = \bar{x}$, because they have to distort their behavior less in period 1. Bad types however would never deviate to $x^u = \bar{x}$, even given the most optimistic beliefs, given how it is derived.

A.1.2 Partially informative equilibria

Although there is an infinite number of partially informative equilibria, the most ‘attractive’ is the one in which there is no party line voting expected of the politicians. Then politicians get the highest utility. If the Intuitive Criterion rules that one out, it automatically rules out equilibria where (and insufficient amount of) party line voting is expected.

Since the topdown equilibrium provides parties with full information about the politician’s type, the value of getting a promotion in that equilibrium is higher than in the partially informative equilibrium. To eliminate the partially informative separating equilibria, I need to show that the good type would gain from deviating to the fully informative separating equilibrium, given the most positive off path beliefs, while the bad type would never deviate to this equilibrium.

Good types want to deviate to the full information equilibrium if the following condition holds, where I impose the most positive off-path beliefs:

$$g + v_L + \delta \left(\rho(\gamma v_N(\hat{\pi}_P = 1) + (1 - \gamma) v_N(\hat{\pi}_P = \pi) + (1 - \rho)(v_L + g)) \right) \leq \\ \bar{x} \cdot u + (1 - \bar{x}) \cdot g + v_L + \delta(\rho(v_N(1) + (1 - \rho)(v_L + g)))$$

Solving this for $v_N(\hat{\pi} = 1)$ gives the how much higher the benefit of being promoted with good beliefs must be

$$v_N(\hat{\pi}_P = 1) \geq \left(\frac{r}{r - 1} \right) \left(v_N(\hat{\pi}_P = \pi) + \frac{1}{r\rho}(1 - \rho)(v_L + r) \right)$$

Given that, good types would deviate to the full information equilibrium.

Second, would bad types want to deviate to this equilibrium? In the partial information equilibrium, they sometimes get promoted or reelected. In the separating equilibrium, they do not, because \bar{x} is sufficiently high that they would not choose to mimic good types.

$$r + v_L + (1 - \gamma)(\rho v_N(\hat{\pi} = \pi) + (1 - \rho)(r + v_L)) \geq r + v_L$$

So bad types would not deviate, even given the positive beliefs. So good types would deviate, given the condition on $v_N(\hat{\pi} = 1)$, while bad types would not.

Intuitively, the intuitive criterion eliminates all separating equilibria if the increase in the value of higher office outweighs the costs for good types. These costs consist of following the party line and the opportunity cost of being promoted with pooling beliefs. If these conditions hold, both parties and politicians are in their most favored equilibrium.

B Extension: Endogenous entry

B.1 Set-up

Players

1. Politicians differ in two dimensions: motivation and ability. In the pool of potential politicians, a mass of π is of the good type, and $1 - \pi$ are bad. Every period, these politicians decide whether or not they want to run for office, which is determined by their ability a_j , which is uniformly distributed between 1 and 2. $a_j \sim U[1, 2]$. A politician with an ability a_j has an outside option of $s \cdot a_j$.

$$U_G(a_j) = v_L + a_j \int_0^1 w(x_{i,1}) di + \delta \left[\left(v_L + a_j \int_0^1 w(x_{i,2}) di \right) \cdot \mathbb{1}_R + \rho v_N \cdot \mathbb{1}_N \right]$$

$$U_C(a_j) = v_L + a_j \int_0^1 r(x_{i,1}) di + \delta \left[\left(v_L + a_j \int_0^1 r(x_{i,2}) di \right) \cdot \mathbb{1}_R + \rho v_N \cdot \mathbb{1}_N \right]$$

2. For simplicity, I assume that ability is not observable by voters nor parties. They do make an inference about the expected ability, given the who decides to run for office, but learn no additional information from the action of the politician.¹⁰
3. Voters' utility is equal to:

$$W = a_1 \int_0^1 w(x_{i,1}) di + \delta a_2 \int_0^1 w(x_{i,2}) di$$

Where a_1 is the ability of the politician in office in Period 1 and a_2 is the ability of the politician in office in Period 2.

4. The party's utility is equal to:

$$U_N(T) = \begin{cases} a & \text{If they promote a politician of type } G \text{ and ability } a. \\ 0 & \text{If they promote a politician of type } B. \end{cases}$$

Timing

Period 1:

1. Politicians choose whether or not to be in the pool for politicians for local office and/or the pool of politicians running for higher office.
2. The rest of the period continues as in the main model, with the only difference being that parties and voters update their beliefs about the distribution of politician's running for office in both periods.

Period 2:

1. Politicians decide whether or not to be in the pool of potential challengers.

¹⁰This prevents the game from turning into a multidimensional signaling model in which politicians signal both their type and ability. Although interesting, this severely complicates the analysis.

2. The rest of the period continues as in the main model.

B.2 Solution

I solve the game for a topdown separating equilibrium in which all good politicians (independent of level of ability) reveal their type by following the party line, while all bad politicians (again, independent of ability), choose to be corrupt.

Selection for period 2

The behavior of good and bad politicians is identical to that in the previous set-up. Given that this is the last period of the game, and there are no more career or re-election incentives, politicians choose whichever policies maximize their utility. These are good policies for the good politicians and bad policies for the bad politicians.

This now becomes relevant, because the policy which politicians choose influence their utility from being in office. That in turn determines whether they want to be in office or not. They only want to run for office if the benefits outweigh their market wage a_j s . That results in two cut-off conditions, for the good and bad types respectively:

$$\begin{aligned} v_L + a_j g &\geq s a_j \\ v_L + a_j r &\geq s a_j \end{aligned}$$

The pool of politicians is then simply equal to the mass of politicians with abilities below this cut-off. This allows me to define cut-off values $a_{G,2}^*$ and $a_{B,2}^*$:

$$\begin{aligned} a_{G,2}^* &= \frac{v_L}{s - g} \\ a_{B,2}^* &= \frac{v_L}{s - r} \end{aligned}$$

This allows me to derive both the probability that a politician in the second period is the good type, as well as their expected ability.

$$\begin{aligned} P(T = G) &= \frac{\pi \cdot (a_{G,2}^* - 1)}{\pi \cdot (a_{G,2}^* - 1) + (1 - \pi) \cdot (a_{B,2}^* - 1)} & E[a_{G,2}^*] &= 1 + \frac{a_{G,2}^*}{2} \\ P(T = B) &= \frac{(1 - \pi) \cdot (a_{B,2}^* - 1)}{\pi \cdot (a_{G,2}^* - 1) + (1 - \pi) \cdot (a_{B,2}^* - 1)} & E[a_{B,2}^*] &= 1 + \frac{a_{B,2}^*}{2} \end{aligned}$$

Period 1

Voters

In the top-down separating equilibrium, voters learn the type of the politician. Furthermore, they can make an inference about the expected types of politicians who would run in period 1, and those who would run in period 2. The question is whether they would still re-elect politicians, given that they learn that the politician is the good type.

It turns out that the answer is yes. For now, suppose that the cut-off value for bad politicians is identical between period 1 and period 2, while the cut-off value for good politicians is higher.

If that is the case, then, conditional on a politician following the party line, they are always re-elected, since if there is positive selection in period 1, (so if $a_{G,1}^* > a_{G,2}^*$), then:

$$\frac{\pi \cdot (a_{G,2}^* - 1)}{\pi \cdot (a_{G,2}^* - 1) + (1 - \pi) \cdot (a_{B,2}^* - 1)} \frac{(2 + a_{G,2}^*)}{2} < \frac{2 + a_{G,1}^*}{2}$$

The first term on the lefthand side is strictly below 1. The second term is smaller than the righthand side, since $a_{G,1}^* > a_{G,2}^*$. The inequality follows.

There are two reasons for positive selection in this equilibrium. First, as I will show below, the cut-off value for bad politicians is identical to the one in derived for period 1. This is because bad politicians reveal their type by not pursuing top down policies, so they do not get re-elected and thus get the utility of being in office for one period. Second, this equilibrium only exists if good politicians prefer revealing their type to pursuing only good policies and not being re-elected. This also implies that the utility of being in office for two periods exceeds the utility from being in office for one period, and thus that the cut-off value is higher.

Parties

For parties, the argument is similar, although it requires an additional assumption.

Since both Good and Bad politicians value higher office equally, their indifference condition and cut-off value are identical.

$$v_N > a_j \cdot s$$

$$a_P^* = \frac{v_N}{s}$$

For parties to promote the good politician that has revealed their type, it is necessary that:

$$E[a|a < a_{G,1}^*] \geq P(T = G|\text{No prior experience}) \cdot E[a|\text{No prior experience}]$$

$$1 + \frac{a_{G,1}^*}{2} \geq \pi \cdot \left(1 + \frac{1}{2} \frac{v_N}{s}\right)$$

$$\pi \leq \frac{a_{G,1}^* + 2}{v_N/s + 2}$$

Where the probability that the politician is the good type on the right-hand side simplifies because both types of politician have the same cut-off condition.

This places a condition on π : learning the type of the politician has to be sufficiently valuable in order to want to promote the politician from local office, rather than someone who selects into ‘just’ running for higher office. It is not necessary that this condition holds. For example, if local office is very unattractive and the probability of promotion is really low, high ability politicians of all stripes rather not run for local office at all. Only if period 1 selection is sufficiently positive (so if $a_{G,1}^*$ is sufficiently large) do parties actually prefer promoting politicians from lower office.

B.3 Politicians' behavior in period 1

The pay-offs of the politician now depend on their ability. This implies politician's willingness to deviate depend on a_j :

$$U_B(x^u = 0, x^g = 0, x^b = 1|a_j) \geq U_B(x^u = \bar{x}, x^g = 0, x^b = 1 - \bar{x}|a_j)$$

$$a_j \cdot r + v_L \geq a_j \cdot (1 - \bar{x}) r + v_L + \delta(1 - \gamma)(\rho v_N + (1 - \rho)(a_j \cdot r + v_L))$$

Solving this for \bar{x} tells us how high \bar{x} has to be for top-down policy making to be a credible signal of type:

$$\bar{x}(\gamma) \geq (1 - \gamma) \frac{\delta}{a_j \cdot r} (\rho v_N + (1 - \rho)(a_j \cdot r + v_L))$$

In a fully separating equilibrium, this has to hold for all a_j , including the one for which it is least likely to hold: $a_j = 1$. Then the condition reduces to the one found in the main text.

Selection for period 1

Because all bad politicians reveal their type, the mass of bad politicians in period 1 and period 2 are equivalent.

$$a_{B,1}^* = a_{B,2}^* = \frac{v_L}{s - r}$$

Good politicians on the other hand do have a different cut-off, as long as they either get re-elected or promoted. This means that there are higher ability politicians running in period 1 than in period 2.

$$U_G = a_j(\bar{x}(\gamma) \cdot u) + (1 - \bar{x}(\gamma)) \cdot g + v_L + \delta(\rho v_N + (1 - \rho)(v_L + a_j g)) \geq a_j s$$

$$a_{G,1}^* = \frac{v_L + \delta(\rho v_N + (1 - \rho)(v_L))}{s + \bar{x}(\gamma)(g - u) - g(1 + \delta(1 - \rho))}$$

This leaves the probabilities of the politician being the good type and their expected ability as:

$$P(T = G) = \frac{\pi \cdot (a_{G,1}^* - 1)}{\pi \cdot (a_{G,1}^* - 1) + (1 - \pi) \cdot (a_{B,1}^* - 1)} \quad E[a_{G,1}^*] = 1 + \frac{a_{G,1}^*}{2}$$

$$P(T = B) = \frac{(1 - \pi) \cdot (a_{B,1}^* - 1)}{\pi \cdot (a_{G,1}^* - 1) + (1 - \pi) \cdot (a_{B,1}^* - 1)} \quad E[a_{B,1}^*] = 1 + \frac{a_{B,1}^*}{2}$$

Since $a_{G,1}^*$ is increasing in γ , which follow from $\bar{x}'(\gamma) < 0$, an increase in the quality of local news (1) increases the probability that the politician is the good type and (2) increases their average expected ability in period 2 is unaffected.

A Proofs for Section 4 (Solution: Topdown separating equilibrium)

Proof. For good politicians, the bad policy is dominated since it delivers no policy pay-off nor positive beliefs. From a policy pay-off perspective, the good policy is preferred to the uniform (since $g > u$), so if they choose u they only set x^u as high as necessary to induce positive beliefs. This is \bar{x} . If they induce negative beliefs, they do not get re-elected or promoted and simply maximize the policy-off by choosing g for all policies.

For bad politicians, corruption rents are linearly increasing in choosing the corrupt policy, so if a strategy induces the belief that a politician is corrupt, a bad politician would maximize their pay-offs by setting $x^b = 1$. There are two ways to induce positive beliefs. One, induce good beliefs, while maximizing corruption rents and hoping to not be uncovered by the newspaper. Since the probability to be discovered does not depend on x^b , setting x^b as high as possible maximizes pay-offs. Or choose to not be corrupt. Since neither u nor g generate corruption rents, all linear combinations of them (meeting the constraint that $x_u \geq \bar{x}$) give the same pay-off for the bad politician. \square