# Optimal Investigative Reporting*

John T. Gasper†

January 20, 2010

**Abstract**

This paper develops a formal model of the investigative journalism process. The choice of what to cover is often thought about in terms of the agenda setting power of the media. While fundamentally important,this research often neglects the internal process by which an outlet chooses what issues to investigate. If media outlets provide any oversight of the political process, then the question of what is investigated is central to their monitoring role. The model I develop in this paper analyzes both single and multiple outlet competition and provides conditions where competition leads to an under supply of investigation.

*I would like to thank Doug Dion, Bill Keech, Sean Gailmard, Scott Gehlbach, John Patty and Maggie Penn for their helpful feedback.

†Postdoctoral Fellow, *Philosophy, Politics & Economics*, 313 Claudia Cohen Hall, 249 South 36th Street, University of Pennsylvania, Philadelphia, PA 19104. http://www.sas.upenn.edu/~jgasper, (215) 746-3618, *jgasper@sas.upenn.edu*

There is a large and growing literature on the effects of the media on voters. Much of this research focuses on the agenda-setting power of the press, but relatively little attention has been given to understanding the foundations of how a news outlet chooses what to cover. This paper is a first step in developing a rigorous investigation into understanding this process. I develop a formal model of one of the keystones of the "fourth estate," the investigative journalism process during a campaign.

Over a quarter of a century ago, the monitoring role of the media reached its pinnacle with Watergate. *The Washington Post*'s reporters Bob Woodward and Carl Bernstein ran a sequence of in depth investigative reports that lead to Congressional investigations, and ultimately with Nixon resigning as President. Since Watergate, however, Bernt and Greenwald (2000) demonstrate that there has been a steady decline in the number investigative reports published by major national newspapers. Moreover, major political scandals are no longer typically being broken by papers such as the *Washington Post* or the *New York Times*, but by outlets such as *The Drudge Report* and *The National Enquirer* initially reporting the Clinton-Lewinsky scandal and John Edward's affair respectively. The results from the model developed here would predict with these empirical regularities.

There are many stories an outlet may cover, but only a limited amount of time or space is available. The choice of what to cover however is often thought about in terms of the press' power to shift public opinion and dictate which stories are "most important." The research on agenda setting, however, often looks at changes in public opinion in response to the media's selection of issues. This research, while fundamentally important, by and large neglects the internal process of how an outlet chooses what issues to initially investigate. As other scholars have noted (McCombs and Shaw 1972; Cohen 1963), the power of the press is not in telling people what to think, but in what to think about. Nonetheless, the analysis of what to initially cover remains relatively unexamined. This paper is a first step toward understanding the foundations of this process. In particular, I analyze the underlying bottom line pressures a news outlet faces in its day to day coverage decisions and begin to address why a news organization would, or would not, ever pursue investigative journalism.

The results of the model focus mainly on the trajectory of coverage leading up to an election. During the campaign an outlet must constantly face the decision between its own investigative reporting versus carrying another "standard" news story. Therefore, over the course of the campaign, for simplicity I assume there are only two stories available for the outlet to cover and that on any given day a news outlet may only cover one of the these.

To highlight the role of investigative reporting, I fix one of the possible stories to be one that comes from a wire service, or the Associated Press (AP). While there is a fair amount of variation in what is covered between major outlets, many news outlets do frequently have a large percentage of stories from the Associated Press (Graber 2006). Therefore it seems reasonable to assume that a story from the wire should provide a baseline for an outside option, against which other stories should be compared.

At a broader level, the model in this paper can be thought of as giving traction to the much broader question of what makes news. I model the news organization as a Bayesian decision maker, and use techniques developed in the theory of optimal sequential decision theory. There is a large literature in psychology on how individuals systematically deviate from Bayesian rationality by various heuristics and biases. Therefore this modeling assumption might seem suspect. It is not my goal, however, to model the exact process by which journalists and editorial boards make their decisions, but to highlight the underlying logic.

# 1 Related Literature

There is a large amount of empirical research, and a growing formal theory literature attempting to understand the role and influence of the media. Recent attention from both the popular and scholarly press has focused on assessing the ideological positions of news outlets. This is a crucial component to understanding the media and information aggregation broadly construed.

In the emerging empirical literature, recent work on measuring the implicit media bias has been accomplished with varying degrees of success. Gentzkow and Shapiro (2006b) examine the heterogeneity of U.S. Daily Newspapers by looking at the partisan correlation between two and three word phrases in the congressional record. Ho and Quinn (2007) provide a novel measurement of ideology using recent statistical classification techniques on expressed political positions. Groseclose and Milyo (2005) estimate media (ADA) ideal points by leveraging citation data of various think tanks from both legislators and media content.

Several recent models of media bias indicate varying theoretical reasons for observing news outlets slanting coverage (Baron 2006; Burke 2008; Gasper 2009; Gentzkow and Shapiro 2006a; Mullainathan and Shleifer 2005). Previous work by Bovitz, Druckman, and Lupia (2002) studies when a news outlet can shape public opinion while dealing with opposing internal conflicts and market forces.

In a slightly different vein, the mechanisms behind information provision have also garnered more attention recently. The model I present in this paper is directly related to this literature and developed from the theory of optimal sequential decision theory. Perhaps oddly, this approach has yet to be extensively applied to the study of the media. However, it is similar to recent work by Patty (2007) on the politics of information gathering in a bureaucracy.

## 2 The Single News Outlet

During the course of a campaign, assume that each day a news outlet may do one of two things. It can either pick up a story from the Associated Press (AP), or it can do its own investigative journalism on some other story or (potentially developing) scandal, $S$. Initially, assume that there are $T$ days until the election. Also assume that every day the outlet faces the same decision, although with potentially different information about $S$, and denote the choice in the $t^{\text{th}}$ stage by $a_t \in A = \{AP, S\}$.

The utility of reporting, or choosing, $a = AP$ gives a fixed known reward, $\lambda \in \mathbb{R}$. However the benefit of covering $S$ is unknown before any investigation takes place. In other words, before any details are gained, the outlet does not know if the story "has legs," and is valuable or if the story is nearly worthless.

Assume that the worth, or impact, of the story is drawn according to a known distribution $F$ with at least one unknown parameter $\omega$. For simplicity, let $\omega$ be the the *unknown* mean of the distribution.[1] In other words, assume that journalists have some understanding about how the world works and that data will be generated according to $f(\cdot|\omega)$, but they do *not* know the true state of the world, $\omega$. Also assume that the journalist has some prior belief

[1]Obviously if the distribution $F$ was a Gamma, or exponential, etc, then this unknown parameter isn't the mean, but determines the mean.

about the underlying state $\omega$, denoted by $\xi$, with mean $\mu \geq 0$. Put differently, the journalist believes there is potentially some worth to the story, but does not know if it is really big or really small.

To be a bit more precise, denote the reward from selected story in stage $t \in \{1, \ldots, T\}$ by a random variable, $Z_t$. Therefore $Z_t = \lambda$ if the paper selects the known option, i.e., $a_t = AP$. Similarly, if the paper chooses to investigate and cover the uncertain story, $a_t = S$, then $Z_t$ is a draw from $F$. The newspaper's strategy, $\sigma$, is simply a mapping from the history of observations to the set of stories $A$. Specifically, $\sigma$ indicates which story to cover in the next stage, given what has been observed, or with $Z_{t-1} = z_{t-1}$, $\sigma_t(z_{t-1}) = a_t$.

If the journalist decides to investigate the scandal, i.e., $a_t = S$, then denote the reward as $X_t$, and assume that conditional on $\omega$, $X_{t'}$ and $X_{t''}$ are independent for all $t'$ and $t''$. The value of $X_t$ is only realized if $a_t = S$. It follows that

$$Z_{t+1} = X_{t+1}\mathbb{I}_{\sigma(z_t)=S} + \lambda\mathbb{I}_{\sigma(z_t)=AP}$$

where $\mathbb{I}$ is the indicator function. The newspaper's objective is therefore to maximize the expected sum of rewards over the $T$ days:

$$\max u(\sigma) = E\left(\sum_{t=1}^{T} Z_t\right).$$

Since there are a finite number of days, the optimal path of investigation leading up to the election, $\sigma^*$, can be found by backward induction and used in the first main result.

**Proposition 1** (Timing of investigation)**.** *If there exists a period $t$ such that $\sigma_t^* = S$, then $\sigma_1^* = S$. In other words, if any investigation takes place, it will happen early.*

Proposition 1 does not say that investigation will not occur after the first stage. It simply says that if any investigation occurs, then it will happen early and specifically in the first stage. If the posterior distribution of $\omega$ is sufficiently high, then continued investigation is optimal. This result depends crucially on the assumption that there is only one news outlet. in section 3 analyzes the situation where there are multiple outlets learning from each other.

The contrapositive of proposition 1, or that if $S$ is not selected in the first stage, then it never will be selected, yields a fair amount of traction. The following section focuses on conditions for selecting $S$ in the first stage. Therefore for the remainder of the paper will restrict attention to a two-stage process ($T = 2$).

## 2.1 *Initial Investigation*

Suppose there are two stages and that in $\sigma^*$ the choice of $AP$ is optimal in the first stage. It follows that the expected benefit over the course of the campaign is simply $2\lambda$. Alternatively, suppose that the optimal strategy indicated that investigation was optimal in the first stage, and denote this by $\sigma^*_{a_1=S}$. The decision at the second stage is between the $AP$ story (yielding a guaranteed payoff of $\lambda$) or a second round of investigation. Realize, however, that there is now a *posterior* distribution about $\omega$. Because some investigation was conducted at the previous stage, there has been information gained. In the second stage, if the posterior distribution of $\omega$ is such that the expected gain from $S$ is greater than $\lambda$, then the optimal strategy is to select $S$.

In the first stage, the expected benefit of investigating $S$ is known from the prior beliefs, $\xi$; denote this expectation by $\mu$. However, this is not the only gain to the journalist by selecting $S$. There information to be learned, and hence there could be value to that information. In the second stage, what is important is that the the expectation of another draw from $F$ is greater than the certain $\lambda$, or $E(X_2|X_1) \geq \lambda$. Therefore the expected gain from the strategy of initially investigating and reporting the potential scandal, $a_1 = S$, is

$$E[u(\sigma^*_{a_1=S})] = \mu + \max\{E(X_2|X_1), \lambda\}.$$

Therefore the optimal strategy for the newspaper is to initially investigate the scandal only if the expected payoff is greater than the safe $AP$ option, or $E[u(\sigma^*_{a_1=S})] \geq 2\lambda$.

Alternatively the above condition can be reformulated such that initial investigation is optimal if

$$\mu - \lambda + \max\{E(X_2|X_1) - \lambda, 0\} \geq 0. \tag{2.1}$$

In the first stage, the newspaper has yet to decide between $S$ and $AP$, and hence has yet to observe $X_1$. There has therefore not been any information gained. However, the $\max\{E(X_2|X_1) - \lambda, 0\}$ term in equation (2.1) is a random variable, and is distributed according to some distribution function $G$.

In the initial decision between $S$ and $AP$, the value of information gained from investigating $S$ is fundamentally important. Clearly, the posterior beliefs about $\omega$ will move with the data that could be gathered. Then before any such data are gathered, there exists a *distribution of the posterior*, calculated by the prior, $\xi$, and the journalist's theory about how the data are generated, the conditional $f(\cdot|\omega)$. From these, one may derive $G$. It then follows that the expectation of the random variable $\max\{E(X_2|X_1) - \lambda, 0\}$ may be expressed:

$$T_G(\lambda) = \int_{\lambda}^{\infty} (x - \lambda) dG(x).$$

The condition for investigating in the first period can now be rewritten as

$$\mu + [T_G(\lambda) - \lambda] \geq 0$$

with $T_G(\lambda)$ having several nice properties:

**Remark 1.** *Given that $G$ is a continuous probability distribution with a finite mean $\vartheta$, $T_G(\lambda)$ is* nonnegative, convex, strictly decreasing,

$$T_G(\lambda) \geq \vartheta - \lambda \text{ for } -\infty < \lambda < \infty,$$
$$\lim_{\lambda \to -\infty} [T_G(\lambda) - (\vartheta - \lambda)] = 0, \text{ and}$$
$$\lim_{\lambda \to \infty} T_G(\lambda) = 0.$$

This formulation of the condition is central to the state the next main result. Informally, it states that there will always exist a cutpoint (in terms of the outside option) that dictates when sampling in the first period is optimal.

**Proposition 2.** *There exists $\lambda_0$, such that for any $\lambda \leq \lambda_0$, initially investigating is always optimal.*

The intuition behind the result is clear. If the outside option is too low, or if the AP news is boring enough, then it will be worthwhile to investigate $S$. In the first stage there is some value of the information gained by investigating. What is perhaps more surprising is that there will always be cases in which sampling is initially optimal, i.e., there *always* exists a $\lambda_0$.

Recall that there have been no specific distributional assumptions made for the above results. In the next subsection, details of the story are assumed to follow a normal distribution, allowing the derivation of various comparative statics.

## 2.2 *Investigating from a Normal Distribution*

In line with the theory developed, continue to assume that $a_t = AP$ gives a fixed known reward, $\lambda$, and that the benefit of covering $S$ will still be unknown for both stages. However, two additional assumptions will be needed:[2]

[2]The results are presented using the precision rather than the variance because it makes them much cleaner. This is completely innocuous because the precision is simply the inverse of the variance.

**Assumption 1.** *The uncertainty about the story, $S$, is specified by a normal distribution with an* unknown *mean, $\omega$, and known precision $r$.*

**Assumption 2.** *The journalist's prior belief about the story, $\xi$, is given by a normal distribution with mean $\mu$ and precision $\tau$.*

This section continues to focus on the two stage setting. The newspaper's objective is, again, to maximize $E(Z_1 + Z_2)$ by choice of $\sigma$. Again, backward induction can be used to solve for $\sigma^*$. If a newspaper initially chooses $a_1 = AP$, then the logic behind proposition 1 applies. There will be no information gained about $\omega$. Hence the optimal action in stage 2 will be $AP$ as well, and the total reward to the newspaper from starting with $a_1 = AP$, will be denoted $u(\sigma^*_{a_1=AP}) = 2\lambda$.

Now suppose that it is optimal to investigate the story in the first stage, i.e., $\sigma^*$ indicated that $a_1 = S$. As described above, the decision at the second stage is between covering the certain ($AP$) story, yielding with a guaranteed payoff of $\lambda$, or to investigate given beliefs from the *posterior* distribution. Because some investigation was conducted in the first stage, there has been information gained about $\omega$. DeGroot (2004) has shown that the posterior of $\xi$ will be a normal distribution with a mean $\frac{\tau\mu+r\bar{X}}{\tau+r}$ and precision $\tau + r$ .

In the second stage, the updated beliefs serve as a best guess of the result from investigating a second time. Therefore the journalist's best guess as to the result of choosing $S$, given he observed $X_1$ in the first period, is simply the mean of the posterior distribution, or

$$E(X_2|X_1) = \frac{\tau\mu + rX_1}{\tau + r}.$$

It follows that the optimal strategy is to investigate a second, and final, time, only if the expected gain is greater than the “safe” option. More specifically,

$$\sigma^*(z_1) = \begin{cases} AP, & \text{if } \frac{\tau\mu + rX_1}{\tau + r} < \lambda; \\ S, & \text{otherwise.} \end{cases}$$

In the first stage, the journalist's initial belief about the newsworthiness of $S$ are given by $\xi$. Therefore the expected reward in the first stage, $E(X_1)$, is equal to $\mu$. Therefore the expected gain from the strategy of initially investigating and reporting the unknown story is

$$E[u(\sigma^*_{a_1=S})] = \mu + \max\{E(\frac{\tau\mu + rX_1}{\tau + r}), \lambda\}.$$

Therefore the optimal strategy for the newspaper is to initially investigate only if the expected payoff is greater than the safe $AP$ option. As it was highlighted in the above section, investigating initially depends on the sign of $\mu - \lambda + \max\{E(\frac{\tau\mu + rX_1}{\tau + r}) - \lambda, 0\}$. If the expression is positive, then investigating in the first stage is optimal. When deciding this, however, $X_1$ has yet to be observed. Therefore the $\max\{E(\frac{\tau\mu + rX_1}{\tau + r}) - \lambda, 0\}$ term is a random variable distributed according to $G$. As before, the expectation of this random variable is:

$$T_G(\lambda) = \int_\lambda^\infty (x - \lambda)dG(x).$$

The calculation of $G$ may be found in the appendix, but the crucial components are stated in the following lemma.

**Lemma 1.** *If assumptions 1 and 2 are met, then $G$ is a normal distribution with mean $\mu$ and precision $\frac{\tau(\tau+r)}{r}$.*

While proposition 2 indicates that there there will exist a $\lambda_0$ that induces a cut point strategy, assumptions 1 and 2 may now be used to derive further results. In particular, the next result indicates the relationship between the optimal initial strategy and the amount of uncertainty contained in the prior belief.

**Proposition 3.** *If assumptions 1 and 2 are met, then as the precision of the prior distribution goes to 0, initially sampling is always optimal.*

Notice that the statement of the proposition is without reference to $\mu$ or $\lambda$. In other words, no matter how great the value of the story is from the known $AP$ option, initially sampling is always optimal if there is enough initial uncertainty. The intuition behind the result comes from the fact that as the dispersion of $\xi$ increases there must be enough mass in the tail of the distribution to outweigh $\lambda$. This also intuitively makes sense: in the first stage there is some value of information to be gained by sampling. Proposition 3 states that, given the distribution assumptions, there will always be cases in which sampling is initially optimal.

# 3 Investigation with Two News Outlets

Now consider a simplified two outlet extension of the model. As before, assume that there are two periods ($T = 2$) and in each period the news outlets may each cover two possible stories, $AP$ or $S$. Again, selecting $AP$ is safe for outlet $i$ and yields a payoff of $\lambda_i$ with certainty. Whereas the payoff from covering $S$ is unknown and depends on an unknown state of the world, $\omega$. It was previously assumed that worth of $S$ followed some unknown Gaussian distribution, but to simplify the analysis in this section assume that $S$ is either newsworthy or not. Therefore assume that there are two states of the world, $\omega \in \{G, B\}$, and when $\omega = G$, $S$ is a "Good" story and yields a payoff $+1$ with probability $q \in (1/2, 1)$ and $-1$ with probability $(1 - q)$. In the "Bad" state of the world ($\omega = B$), the story yields $-1$ with probability $q \in (1/2, 1)$ and $+1$ with probability $(1 - q)$.

Denote the outlets' common prior belief that $S$ is a good story by $p_0$. Furthermore, assume that all actions and stage payoffs are publicly observed. Hence let $p_t$ be the posterior of this belief in period $t$ (that is, updating on all information in periods $1, \ldots, t-1$). Notice that in any period $t$ the expected payoff of investigating the risky story, $S$, is $p_t(2q-1) + (1-p_t)(1-2q) = (2p_t - 1)(2q-1)$.

Suppose the two outlets are different in that they receive different payoffs from selecting the "safe" story. Specifically, payoffs $\lambda_i$ are received with when $AP$ is selected by outlet $i$. Also assume that outlet one is more reputable, or "advantaged," by letting $Y_1 > Y_2$.[3] Therefore the two-by-two normal form representation of this setting is shown in Table 1[4]

| | | Outlet $j$: $S$ | Outlet $j$: $AP$ |
|---|---|---|---|
| Outlet $i$ | $S$ | $E[X_i^1] + E[\max(E[X_i^2 \mid X_i^1, X_j^1], \lambda_i)]$ | $E[X_i^1] + E[\max(E[X_i^2 \mid X_i^1], \lambda_i)]$ |
| | $AP$ | $\lambda_i + E[\max(E[X_i^2 \mid X_j^1], \lambda_i)]$ | $\lambda_i + E[\max(E[X_i^2], \lambda_i)]$ |

Table 1: Normal form representation of the game in the first period. Note that the game's structure is symmetric but only Outlet $i$'s payoffs are given.

Then outlet 1's decision is between initially selecting $S$, yielding an immediate expected payoff of $E(X_1^1)$, or just selecting the $AP$, yielding a "safe" payoff of $Y_1$. Yet, this is not the only gain to outlet 1. There is information to be learned about the underlying state of the world; i.e., whether $S$ is a good story or not. Hence, there is some value to that information.

If in the first period both outlets selected $S$, then the expected payoff from $S$ in the second period is $E(X_1^2 \mid x_1^1, x_2^1)$. Whereas if only outlet 1 selects $S$ in the first period, only that observation would inform both outlets about expected payoffs in period 2, $E(X_1^2 \mid x_1^1)$. The value to outlet $i$ by initially investigating $S$ can therefore be thought of in terms of the

[3] The symmetric case, $\lambda_1 = \lambda_2$, is available upon request.

[4] Note the slight change in notation: $X_i^k$ stochastic payoff to outlet $i$ in period $k$, with the lowercase $x_i^k$ denoting the realized payoff to $i$ in period $k$.

expected payoff today plus the expected payoff from optimal decision tomorrow, *given the information learned.*

$$E[X_i^1] + E\left[\mathbb{I}_J(S)\max[E(X_i^2|X_i^1, X_j^1), \lambda_i] \quad + \quad \mathbb{I}_j(AP)\max E(X_i^2|X_i^1), \lambda_i]\right] \tag{3.1}$$

Whereas initially investigating $AP$ for outlet $i$ yields:

$$\lambda_i + E\left[\mathbb{I}_j(S)\max[E(X_i^2|X_j^1), \lambda_i] \quad + \quad \mathbb{I}_j(AP)\max[E[X_i^2], \lambda_i]\right] \tag{3.2}$$

for $i, j \in \{1, 2\}$ and $i \neq j$.

Notice that unlike the previous section, initially selecting $AP$ is no longer irreversible along the optimal path and Proposition 1 does not hold. Previously, if an outlet ever optimally chose the $AP$ story, learning about the unknown state would stop. Therefore, it would then also be optimal to select the $AP$ option in every future period. The interactive search problem is different.[5] In particular, there is an opportunity for an outlet to learn about the state of the world from the other outlet.

Since $\lambda_1, \lambda_2$ and $q$ are parameters of the model, it is possible to solve for the minimum belief needed for an outlet to initially select story $S$. To highlight the difference between the single versus dual decision maker case, define two functions that specify these cut-points for outlet $i$'s initial belief. Let the point $p^*(Y, q)$ is the lowest value of of the initial beliefs for which a *single outlet* would be willing to investigate story $S$ in the first period. Whereas the cut-point defined by $\bar{p}_i(\lambda_1, \lambda_2, q)$ is the minimal prior belief needed for outlet $i$ to initially select $S$ in the *two outlet* case. Similarly, the set of beliefs less than or equal to $\underline{p}_i(\lambda_1, \lambda_2, q)$ defines the region where outlet $i$ taking action $AP$ is optimal with two outlets.

[5]While this is technically possible in the symmetric two candidate case ($\lambda_1 = \lambda_2$), I avoid the presentation here because it can only occur when outlets are both mixing between the risky and safe options.

In the current setting, outlet 2 will initially raise $S$ only if the benefit of doing so is greater than not investigating the story: that is, when equation 3.1 is greater than equation 3.2, for $i = 2$. Setting these equations equal and solving for $p$ will yield a $\bar{p}_2(\lambda_1, \lambda_2, q)$, such that any for any $p > \bar{p}_2(\lambda_1, \lambda_2, q)$ outlet 2 will find it strictly dominant to investigate $S$ in the first period.

Since $\underline{\mathrm{p}}_i(\lambda_1, \lambda_2, q)$ and $\bar{p}_i(\lambda_1, \lambda_2, q)$ are increasing in $\lambda_i$ and by assumption $\lambda_1 > \lambda_2$, it follows that $\underline{\mathrm{p}}_2(\lambda_1, \lambda_2, q) < \underline{\mathrm{p}}_1(\lambda_1, \lambda_2, q)$ and $\bar{p}_2(\lambda_1, \lambda_2, q) < \bar{p}_1(\lambda_1, \lambda_2, q)$. Therefore there are two cases; either $\underline{\mathrm{p}}_1(\lambda_1, \lambda_2, q) \geq \bar{p}_2(\lambda_1, \lambda_2, q)$ or it is not.

*Case 1:* Let $\underline{\mathrm{p}}_1(\lambda_1, \lambda_2, q) \geq \bar{p}_2(\lambda_1, \lambda_2, q)$. In this case, both outlets know that outlet 1 will not investigate the risky story when when priors are below $\underline{\mathrm{p}}_1(\lambda_1, \lambda_2, q)$. Investigating story $S$ in this range of beliefs is dominated for outlet 1. Therefore, $\mathbb{I}_1(S) = 0$ and outlet 2 acts as if it were a single decision maker and the cut-point for initially selecting $S$ is $p^*(\lambda_2, q)$. Outlet 2's problem then reduces to finding a value of $p$ such that:

$$E(X_2^1) + E(\max\{E(X_2^2|X_2^1), \lambda_2\}) = 2\lambda_2. \tag{3.3}$$

As a result, $\underline{\mathrm{p}}_2(\lambda_2, q) = \bar{p}_2(\lambda_2, q) = p^*(\lambda_2, q)$. Then for any $p > \underline{\mathrm{p}}_2(\lambda_1, \lambda_2, q)$, it is dominant for outlet 2 to investigate $S$ in the first period.

Consider $p \in (\underline{\mathrm{p}}_1(\lambda_1, \lambda_2, q), \bar{p}_1(\lambda_1, \lambda_2, q))$. Outlet 2 researches $S$ in in this range (since by assumption $\underline{\mathrm{p}}_1(\lambda_1, \lambda_2, q) > \bar{p}_2(\lambda_1, \lambda_2, q)$) and outlet 1 best-replies. Since $\lambda_1 > \lambda_2$, outlet 1 knows that outlet 2 will sample and hence 1 knows the it will obtain some "free" information. Therefore solving for $\bar{p}_1(\lambda_1, \lambda_2, q)$, outlet 1 solves the following:

$$E(X_1^1) + E(\max\{E(X_2^1|X_1^1, X_1^2), \lambda_1\}) = \lambda_1 + E(\max\{E(X_2^1|X_1^2), \lambda_1\}) \tag{3.4}$$

When $p > \bar{p}_1(\lambda_1, \lambda_2, q)$ (and hence $p > \underline{\mathrm{p}}_2(\lambda_1, \lambda_2, q)$) both players will investigate story

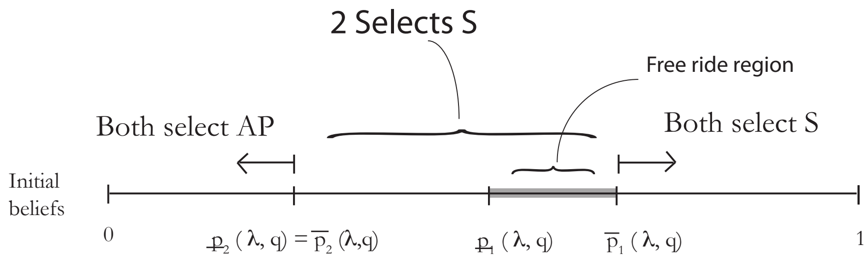


Figure 1: Minimum beliefs needed to sample $S$

$S$ in the first round, because both believe that it is likely that $S$ is a good story and will outweigh the gain from $AP$.

Whereas when $p > \underline{p}_2(\lambda_1, \lambda_2, q)$ but $p < \bar{p}_1(\lambda_1, \lambda_2, q)$ only player 2 will initially cover story $S$. In this second case, outlet 2 can be seen as gambling with a controversial story. Recall that outlet 1's assumed (reputation) advantage has led to $\lambda_1$ being greater than $\lambda_2$. The new, or less reputable outlet, finds itself in a position where it must gamble.

Next consider the region when $p < \bar{p}_1(\lambda_1, \lambda_2, q)$ but $p > \underline{p}_1(\lambda_1, q)$, which is denoted as the "Free ride region" in Figure 1. In this region, outlet 2 again investigates the risky story and 1 does not, but if outlet 1 were alone then it *would have investigated.* In other words, outlet 1 believes ex ante that $S$ is a good story, but can rely on outlet 2 to take the risk of researching it since 2 has a strictly dominant strategy to sample. At some level, the advantaged outlet can be seen as exploiting the risk of sampling from outlet 2.

*Case2:* Alternatively it is possible that $\underline{p}_1(\lambda_1, \lambda_2, q) < \bar{p}_2(\lambda_1, \lambda_2, q)$. In the Bernoulli case developed in this section, this occurs precisely when $\lambda_1 < \frac{4q^2-4q+3\lambda_2+1}{3+\lambda_2}$. In this case, found in Figure 2, the advantaged outlet can no longer rely on the same behavior from outlet 2. Interestingly, the behavior is very similar, but now there is an intermediate region (region C) where the only equilibrium is for both outlets to mix between $AP$ and $S$. In regions A and B in Figure 2, outlet 2 still has a strictly dominant strategy to investigate story $S$ and hence the advantaged outlet will only select $AP$.

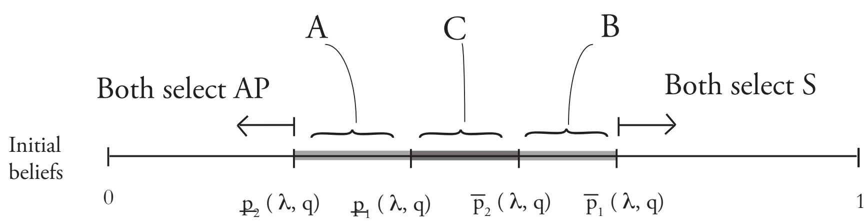


Figure 2: Minimum beliefs needed to initially sample $S$

In general, for each $\lambda_i$ and $q$ there is going to exist a minimum level of belief that the $S$ story is a "good" story. The surface generated by the point of indifference for outlet 2, $E(X_1^2) + E(\max\{E(X_2^2|X_1^2), \lambda_2\}) = 2\lambda_2$ yields exactly this $\underline{p}_2(\lambda_1, \lambda_2, q)$. Any $(p, q, \lambda_1, \lambda_2)$ point below the $\underline{p}_2(\lambda_1, \lambda_2, q)$ surface will result in the disadvantaged outlet selecting the "safe" story.

# 4 Discussion

In terms of American politics and journalism, Watergate was a defining moment. In the aftermath of the popularity Woodward and Bernstein, the interest in becoming an investigative reporter surged yet as Aucoin (2005) notes, there was a steady decline in the number of investigative reports published by major papers. During the same period, profit margins became thinner and thinner and many newspapers adopted a more business like approach to their selection of news stories.

The model presented attempts to highlight these incentives underlying the investigative journalism process. I have assumed that an outlet may only report on one of two options: investigating and reporting the details of a scandal, or picking up a story from the AP. While this is clearly a simplification, considering that many news outlets do frequently place stories from the Associated Press it seems like a fair place to start. The selection of the AP story gives a fixed, known, reward making the model is similar to what is known in the decision theoretic literature as a one-armed bandit.

Alternatively, one could assume that there is some variance in newsworthiness of the AP story. Additionally, one could assume that there are multiple stories an outlet could cover. These extensions would take the model into the realm of what is known as a multi-arm bandit. Mathematically these become more cumbersome but much of the underlying logic would likely remain. With multiple stories, it is likely that the outlet would select the story that gives the highest expected reward at that stage. In particular, I conjecture that there would likely exist an ordering of the stories according to updated beliefs, and selection would take place over the first $k$ of those stories.[6]

The logic of the model is driven by two fundamental forces: the value of information gained by investigating and the opportunity cost involved in investigating. Anecdotal ev-

[6]The mathematical development of sequential decision theory becomes unwieldy very quickly. However, one of the seminal results in its development is is known as the Gittins index. This ordering would likely be given by that index.

idence suggests that no major news outlet is a complete replication of Associated Press. There are journalists with their "normal beats." Different papers are not mirror images of one another and even similar stories have subtle differences.

The model presented in this paper highlights the economic forces that have arguably driven to a reduction of investigative reports. The model also indicates a reason why one would expect to see more risky or scandalous stories broken by tabloids than major news outlets: tabloids have a lower outside option. They simply can not gain market share by reporting the standard pieces that major new outlets pick up. Matt Drudge was not the first to find out about the Clinton-Lewinsky affair; *Newsweek* had been sitting on it until it had stronger verification (Bernt and Greenwald 2000). Newsweek had relatively higher outside options and chose not to gamble with a weakly confirmed report. Therefore, tabloids arguably face a lower downside risk (perhaps in terms of reputation or legal costs) of being wrong. Hence the model would predict that they will run more fringe or riskier investigative pieces.

The model presented can be thought of as an initial step toward providing a rigorous answer to the question of when a news organization would begin to pursue investigative journalism. Relatively little attention has been given to understanding the calculus underlying how a news outlet chooses what to cover. There are many possible stories, but limited space and time. This paper has developed a rigorous investigation into understanding one aspect of this choice.

# Appendix: Proofs

**Proposition 1** *If there exists a period $i$ such that $\sigma_i^* = S$, then $\sigma_1^* = S$. In other words, if any investigation takes place, it will happen early.*

*Proof.* The logic behind this result is straightforward. Since $\sigma^*$ is optimal it must be optimal at every stage. So now suppose there is a stage $j$ in which the optimal choice is to select $AP$. Then no information is gained about $\omega$. Hence the beliefs about $\omega$ are the same in the $j+1$ stage, and the optimal choice in $j+1$ is to select $AP$. In other words, once $AP$ is selected, it will always be selected. It follows that if $S$ is selected, then it will be selected initially. □

**Proposition 2** *There exists $\lambda_0$, such that for any $\lambda \leq \lambda_0$, initially investigating is always optimal.*

*Proof.* From Remark 1, it immediately follows there there exists a fixed point in the $T_G$ mapping. Let $\lambda_0$ be that fixed point, i.e., $T_G(\lambda_0) = \lambda_0$. By the fact that $T_G(\cdot)$ is strictly decreasing, any $\lambda \leq \lambda_0$ will imply that $T_G(\lambda) - \lambda \geq 0$. Hence $\mu + [T_G(\lambda) - \lambda] \geq 0$ as required. □

**Lemma 1** *If assumptions 1 and 2 are met, then $G$ is a normal distribution with mean $\mu$ and precision $\frac{\tau(\tau+r)}{r}$.*

*Proof.* To calculate $G$, first calculate the distribution of $X$ given the prior $\xi$.

By an initial assumption, the conditional distribution of $X$ when $\omega = \bar{\omega}$ is a normal distribution with mean $\bar{\omega}$ and precision $r$. It follows that *ex ante* marginal distribution of $X$ is then a normal distribution with mean $\mu$. One must, however, calculate the precision of outlets ex ante belief about $X$. The precision of $X$ will be the sum of the precision of the conditional distribution, $r$, and the precision of the prior, $\tau$.[7] Since the precision is the inverse of the variance, and hence the precision of the marginal distribution of $X$ is $r + \tau = \frac{r\tau}{\tau+r}$. It follows that $G$ is a normal distribution with mean $\mu$, and the precision of $G$ is $\frac{\tau(\tau+r)}{r}$ □

**Proposition 3** *If assumptions 1 and 2 are met, then as the precision of the prior distribution goes to 0, initially sampling is always optimal.*

[7] DeGroot (2004), p 263.

*Proof.* By lemma 1 if follows that $G$ is a normal distribution, and for notational simplicity let us refer to it's precision as $\nu$. DeGroot has shown that since $G$ is a normal distribution, $T_G(\cdot)$ has a very elegant property (DeGroot 2004). In particular, $T_G(\cdot)$ can be expressed in terms of a transformation of $\Phi$, the standard normal distribution (mean=0 and precision=1):

$$\begin{aligned} T_G(\lambda) &= \nu^{-\frac{1}{2}}\Psi[\nu^{\frac{1}{2}}(\lambda-\mu)] \\ \text{where } \Psi(s) &= \int_s^\infty (x-s)d\Phi \end{aligned}$$

The condition for initially investigating may now be written $\nu^{-\frac{1}{2}}(\Psi[\nu^{\frac{1}{2}}(\lambda-\mu)]-\nu^{\frac{1}{2}}(\lambda-\mu))\geq 0$. Since $\nu>0$, this amounts to $\Psi[t]-t\geq 0$ where $t=\nu^{\frac{1}{2}}(\lambda-\mu)$.

By proposition 2 must exist a fixed point $\Psi(t^*)=t^*$. For $t>t^*$, the newspaper will always go with the safe option, i.e., cover the $AP$ story. For $t\leq t^*$, initially investigating is optimal. Observe that it is required that $\sqrt{\frac{\tau(\tau+r)}{r}}(\lambda-\mu)\leq t^*$, and that $\frac{\tau(\tau+r)}{r}$ goes to 0 as $\tau\to 0$. In other words, if the journalists initial prior uncertainty is large enough, then this condition will always be met. □

# References


James L. Aucoin. *The Evolution of American Investigative Journalism.* University of Missouri Press, 2005.

David Austen-Smith and Arthur Lupia. Information in elections. In John H. Aldrich, James E. Alt, and Arthur Lupia, editors, *A Positive Change in Political Science: The Legacy of Richard D. McKelvey's Most Influential Writings.* University of Michigan Press, Forthcoming.

David P. Baron. Persistent media bias. *Journal of Public Economics*, 90(1-2):1–36, January 2006.

James O. Berger. *Statistical Decision Theory and Bayesian Analysis.* Springer, 1985.

Joseph Bernt and Marilyn Greenwald. *The Big Chill: Investigative Reporting in the Current Media Enviroment.* Iowa State University Press, 2000.

Gregory Bovitz, James Druckman, and Arthur Lupia. When can a news organization lead public opinion? *Public Choice*, 113:127ñ155, 2002.

Jeremy Burke. Primetime spin: Media bias and belief confirming information. *Journal of Economics and Management Strategy*, 17(3):633–665, 2008.

Randall L. Calvert. The value of biased information: A rational choice model of political advice. *Journal of Politics*, 47:530–555, May 1985.

Bernard Cohen. *The Press and Foreign Policy*. Princeton University Press, 1963.

Morris DeGroot. *Optimal Statistical Decisions*. Wiley Classics Library, 2004.

Baruch Fischhoff and Ruth Beyth-Marom. Hypothesis evaluation from a bayesian perspective. *Psychological Review*, 90(3):239–260, 1983.

Sean Gailmard and John W. Patty. Zlackers and zealots: Civil service, policy discretion, and bureaucratic expertise. *American Journal of Political Science*, 51(4):873–889, October 2007.

John T. Gasper. Reporting for sale: the market for coverage. *Public Choice*, 141(3), 2009.

Matthew Gentzkow and Jesse M. Shapiro. Media bias and reputation. *Journal of Political Economy*, 114(2):280–316, 2006a.

Matthew Gentzkow and Jesse M. Shapiro. What drives media slant? Evidence from U.S. daily newspapers. *NBER working paper series*, (w12707), 2006b.

Lisa George and Joel Waldfogel. Who affects whom in daily newspaper markets? *Journal of Political Economy*, 11:765–85, 2003.

Alan Gerber, Dean Karlan, and Daniel Bergan. Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. Mimeo, Yale University, April 2007.

Doris A. Graber. *Mass Media and American Politics*. CQ Press, 2006.

Tim Groseclose and Jeff Milyo. A measure of media bias. *Quarterly Journal of Economics*, 120(4):1191–1237, 2005.

Daniel E. Ho and Kevin M. Quinn. Assessing political positions of media. Working paper, 2007.

Valentino Larcinese. Information acquisition, ideology and turnout: Theory and evidence from britain. *Journal of Theoretical Politics*, 21(2):237–276, April 2009.

Maxwell E. McCombs and Donald L. Shaw. The agenda-setting function of mass media. *The Public Opinion Quarterly*, 36(2):176–187, 1972.

Sendhil Mullainathan and Andrei Shleifer. The market for news. *American Economic Review*, 95(4):1031–1053, September 2005.

John W. Patty. Top down bias. Harvard working paper, 2007.

Riccardo Puglisi and James Snyder. Media coverage of political scandals. Unpublished Working Paper. SSRN 1117716, April 2008.

Joel Waldfogel and Lisa George. The new york times and the market for local newspapers. *American Economic Review*, 2006.